

شمارش، حساب، اندازه‌گیری: نخستین‌هایی که آدمیان از دیرباز تا امروز آموخته‌اند تا برای ورود به رقابتی بی‌وقفه آماده شوند: آدمیانی که با انبوهی از معیارها روبه‌رو می‌شوند و خواسته و ناخواسته، آگاهانه و ناآگاهانه در تلاشی همیشگی به دنبال هرچه نزدیک‌تر کردن «اندازه‌هایشان» با این معیارها هستند. جری زد مولر در کتاب «استبداد معیارها» مخاطب را درباره این معیارها که تقریباً در همه جا رسوخ کرده‌اند آگاه می‌کند و موثر بودنشان را به چالش می‌کشد. او این پرسش را مطرح می‌کند که آیا معیارها همیشه و همه‌جا مفیدند و موجب افزایش کارایی و بهره‌وری و در نهایت رشد می‌شوند؟ آیا «اندازه‌گیری» تنها روش درست است و آیا این روش همواره پاسخگو است؟

مولر اصل معیارها و روش اندازه‌گیری را زیر سوال نمی‌برد و بر مفید بودن آنها به ویژه در برخی حوزه‌هایی که کمتر با انسان سروکار دارند صحه می‌گذارد، اما هوشمندانه نکته‌ای را یادآور می‌شود: در جامعه با انسان روبه‌رو هستیم: انسانی که با انسان دیگر تفاوت دارد، انسانی که می‌تواند نسبت به معیارها واکنش نشان دهد و یا بر اندازه‌گیری آنها اثرگذار باشد، انسانی که تمام ابعادش قابل اندازه‌گیری نیست و انسانی که مجهز به قوه ادراک و قضاوت است.

استبداد معیارها

نویسنده: جری زد مولر

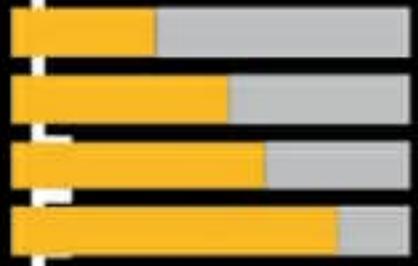
مترجم: زهرا تدین



انجمن کتابخانه‌های ملی و اسناد و کتابخانه‌های ایران  
مرکز نشر



# استبداد معیارها



جرى زد.مولر

مترجم:  
زهرا تدین







# استبداد معیارها

جری زد.مولر

مترجم: زهرا تدین

سرشناسه	: زدمولر، جری. Z.Muller, Jerry
عنوان و نام پدیدآور	: استبداد معیارها / نویسنده: جری زدمولر، مترجم: زهرا تدین
مشخصات نشر	: تهران: سبزان، ۱۳۹۹
مشخصات ظاهری	: ۲۱۶ ص.
شابک	: ۹۷۸-۶۰۰-۱۱۷-۵۸۹-۳
وضعیت فهرست‌نویسی	: فیپا
یادداشت	: عنوان اصلی THE TYRANNY OF METRICS
موضوع	: اندازه‌گیری
موضوع	: استدلال
موضوع	: کسب‌وکار
شناسه افزوده	: تدین، زهرا ۱۳۶۶ - مترجم
شناسه افزوده	: شافعی، غلامحسین ۱۳۳۰ - مقدمه
رده‌بندی کنگره	: ۱۷۵tk/۳۲/ش۹ی۲ ۱۳۹۹
رده‌بندی دیویی	: ۳۸۱/۰۰۱
شماره کتابشناسی ملی	: ۵۰۳۴۷۳۶

## استبداد معیارها

نویسنده: جری زد.مولر

مترجم: زهرا تدین

کاری از: اتاق بازرگانی، صنایع، معادن و کشاورزی ایران، مرکز پژوهش‌ها

مقدمه: غلامحسین شافعی

مدیر طرح: حسین سلاح‌ورزی

ناظر: سجاد نجفی

قیمت: ۷۰۰۰۰۰ ریال

تیراژ: ۱۰۰۰ نسخه

نوبت چاپ: اول، زمستان ۱۳۹۹

انتشارات: سبزان

چاپ و صحافی: کامیاب

شابک: ۹۷۸ - ۶۰۰ - ۱۱۷ - ۵۸۹ - ۳



اتاق بازرگانی، صنایع، معادن و کشاورزی ایران  
مرکز پژوهش‌ها

## سخن آغازین

شمارش، حساب، اندازه‌گیری؛ نخستین‌هایی که آدمیان از دیرباز تا امروز آموخته‌اند تا برای ورود به رقابتی بی‌وقفه آماده شوند. آدمیانی که در خانه، مدرسه، دانشگاه و محیط کار با انبوهی از معیارها روبه‌رو می‌شوند و خواسته و ناخواسته، آگاهانه و ناآگاهانه در تلاشی همیشگی به دنبال هرچه نزدیک‌تر کردن «اندازه‌هایشان» با این معیارها هستند؛ که اگر در این تلاش پیروز شوند به احتمال زیاد پاداش می‌گیرند و اگر مغلوب شوند، احتمالا مجازات و سرخورده خواهند شد. جری زد مولر در کتاب «استبداد معیارها» مخاطب را درباره این معیارها که تقریبا در همه جا رسوخ کرده‌اند آگاه می‌کند و موثر بودنشان را به چالش می‌کشد. او این پرسش را مطرح می‌کند که آیا معیارها همیشه و همه‌جا مفیدند و موجب افزایش کارایی و بهره‌وری و در نهایت رشد می‌شوند؟ آیا «اندازه‌گیری» تنها روش درست است و آیا این روش همواره پاسخگو است؟

مولر اصل معیارها و روش اندازه‌گیری را زیر سوال نمی‌برد و بر مفید بودن آنها به ویژه در برخی حوزه‌هایی که کمتر با انسان سروکار دارند صحه می‌گذارد، اما هوشمندانه نکته‌ای را یادآور می‌شود: در جامعه با انسان روبه‌رو هستیم؛ انسانی که با انسان دیگر تفاوت دارد، انسانی که می‌تواند نسبت به معیارها واکنش نشان دهد و یا بر اندازه‌گیری آنها اثرگذار باشد، انسانی که تمام ابعادش قابل اندازه‌گیری نیست و انسانی که مجهز به قوه ادراک و قضاوت است.

به نظر می‌رسد تنها یکی از این موارد برای بازنگری درباره به کارگیری معیارها و اندازه‌گیری در علوم انسانی و مصادیق عملی آن در بخش‌های مختلف جامعه کافی باشد. این کتاب برای لحظه‌ای توقف، درنگ و تأمل است؛ تأمل نه برای بازآفرینی آن‌چه که تاکنون اندازه‌گیری شده است، بلکه برای بازاندیشی درباره آن‌چه که تاکنون اندازه‌گیری نشده است و در نهایت اندیشیدن به این‌که نه هر آن‌چه که قابل اندازه‌گیری است اهمیت دارد و نه هر آن‌چه که اهمیت دارد قابل اندازه‌گیری است.

خوانش این کتاب به صاحبان کسب‌وکار و مدیران بنگاه‌ها و کارکنان کمک می‌کند تا بدانند محاسن اندازه‌گیری عملکرد در چیست و معایبش در کجاست. در عین حال، این کتاب به طراحان معیارهای اندازه‌گیری عملکرد در بنگاه‌های تولیدی و خدماتی نیز کمک می‌کند تا معیارهای کاراتری را برای اندازه‌گیری عملکرد در سطوح مختلف وظایف بنگاه‌شان به کار بگیرند.

**غلامحسین شافعی**

**ریاست اتاق بازرگانی، صنایع، معادن و کشاورزی ایران**

نویسنده: برای جمیز دوئرتی



«آنهایی که باور دارند آنچه را که نمی‌توانید اندازه‌گیری کنید وجود ندارد، همان‌هایی هستند که بر این باورند آنچه را که می‌توانید اندازه‌گیری کنید وجود دارد.»

آرون هاسپیل



## فهرست مطالب

پیشگفتار.....	۱۳
سپاسگزاری.....	۲۸

### بخش اول: استدلال

فصل ۱: استدلال در چند کلمه.....	۳۳
فصل ۲: اشکالات تکراری.....	۳۹

### بخش دوم: پیشینه

فصل ۳: خاستگاه‌های اندازه‌گیری و پرداخت مبتنی بر عملکرد.....	۴۵
فصل ۴: چرا معیارها بسیار محبوب شدند.....	۵۴
فصل ۵: کارفرمایان، کارگزاران و انگیزه‌ها.....	۶۴
فصل ۶: انتقادات فلسفی.....	۷۳

### بخش سوم: اندازه‌گیری نادرست همه چیز؟ مطالعات موردی

فصل ۷: کالج‌ها و دانشگاه‌ها.....	۸۱
فصل ۸: مدارس.....	۱۰۴
فصل ۹: پزشکی.....	۱۱۷
فصل ۱۰: پلیس.....	۱۳۹
فصل ۱۱: ارتش.....	۱۴۵
فصل ۱۲: کسب‌وکار و مالیه.....	۱۵۱
فصل ۱۳: بشردوستی و کمک خارجی.....	۱۶۷

### ضمیمه

فصل ۱۴: هنگامی که شفافیت دشمن عملکرد است: سیاست، دیپلماسی، اطلاعات و ازدواج.....	۱۷۳
--	-----

### نتیجه‌گیری

فصل ۱۵: پیامدهای منفی ناخواسته اما قابل پیشبینی.....	۱۸۱
فصل ۱۶: چه زمانی و چگونه از معیارها استفاده شود؟.....	۱۸۷

یادداشت‌ها.....	۱۹۷
-----------------	-----



## پیشگفتار

سریال وایر یا شنود<sup>۱</sup> که از شبکه اچ بی او<sup>۲</sup> پخش می‌شد و بر اساس تجارب زندگی واقعی سازندگان آن، دیوید سیمون<sup>۳</sup> و اد برنز<sup>۴</sup>، ساخته شده بود، از دید برخی با دلایل قانع‌کننده به عنوان یکی از بزرگ‌ترین مستندات فرهنگی عصر ما مورد توجه قرار گرفته است. تمرکز سریال تنها بر یک شهر آمریکایی، بالتیمور است و تعداد کمی از نهادهای مهم را واکاوی می‌کند- پلیس، سیستم آموزشی، سیاست شهرداری، رسانه- و تصویری همانند اشعه ایکس را از کارکرد و نقایص آنها ارائه می‌دهد. این سریال برای مخاطبان بین‌المللی نیز جذاب بود زیرا محتوا و مضامین آن درباره اختلالات سازمانی به طور گسترده‌ای در سرتاسر جوامع غربی طنین‌انداز شده است.

یکی از مضامین تکراری در سریال شنود برجستگی معیارهاست: عملکرد اندازه‌گیری شده به‌منزله مشخصه «پاسخگویی». فرماندهان پلیس روی اعداد و افزایش آنها حساسیت و وسواس زیادی دارند-مثلاً، موارد حل‌شده، گرفتن مواد مخدر، آمار جنایت- و با ابزارهای متعددی که

---

1. The Wire

2. Home Box Office

یکی از بزرگ‌ترین پخش‌کننده‌های شبکه‌های تلویزیونی کابلی که زیرمجموعه شرکت برادران وارنر در آمریکا و وابسته به باکس آفیس خانگی است. (ویکی‌پدیای فارسی). م

3. David Simon

4. Ed Burns

کارایی را قربانی می‌کند، به دنبال اهداف آماری هستند. سیاستمداران به دنبال اعدادند تا گواهی باشد بر موفقیت پلیس در کنترل جرم و جنایت. از این‌رو، واحدهای پلیس تمام تلاش خود را می‌کنند تا در منطقه آنها قتل و کشتار اتفاق نیفتد: هنگامی که معلوم می‌شود یک دارودسته مواد مخدر اجساد را در خانه‌های متروکه دفن کرده‌اند، گروه‌بان مسئول این پرونده قتل از این کشف دلسرد می‌شود زیرا موجب کاهش «ترخ پاکسازی»<sup>۱</sup> شده است؛ معیاری که نشان‌دهنده درصد جرایم حل‌شده است. بخش عمده طرح این سریال پیرامون کارآگاهان متخصصی می‌چرخد که به دنبال اقامه یک دعوای جنایی پیچیده در مواجهه با یک ارباب بزرگ مواد مخدر هستند. اما از آنجا که برپا کردن چنین پرونده‌ای، ماه‌ها و حتی سال‌ها به طول می‌انجامد، از انجام این کار زیر نظر مقامات مافوق منصرف می‌شوند زیرا این مقامات از نیروهای پلیس می‌خواهند که معامله‌گران مواد مخدر در سطوح پایین - که تعدادشان بسیار زیاد است - دستگیر کنند که این اقدام موجب بالا رفتن معیارهای مطلوب و مناسب می‌شود؛ البته به‌رغم این واقعیت که افرادی که دستگیر می‌شوند به سرعت با تعدادی دیگر جایگزین خواهند شد. دفتر شهردار تقاضا دارد که میزان جرایم بزرگ پیش از پایان سال ۵ درصد کاهش یابد، هدفی که فقط می‌تواند با چشم‌پوشی از جرایم واقعی و کاهش میزان جدی‌بودن آنها به آن دست یافت. در هر مورد، آنها درگیر «شوخی آماری»<sup>۲</sup> هستند و شاخص‌های آماری خود را با تحریف نتایج واقعی یا منحرف ساختن زمان و تلاش خود برای پیشگیری از جرم و جنایت به سمت کارهای با اثربخشی کمتر بهبود می‌بخشند.

جزء دیگری از داستان این سریال شامل یک نیروی سابق پلیس است که در مدرسه راهنمایی محله‌ای گرفتار فقر، سوءمصرف مواد مخدر و فروپاشی خانواده، تدریس می‌کند. دانش‌آموزان در مدرسه عملکرد

1. Clearance Rate

میزان جرایم و جنایات حل‌شده به کل جرایم و جنایات روی داده. م

2. Juking the stats

ضعیفی دارند و اگر نمرات امتحانی دانش‌آموزان بهبود نیابد، مدرسه در معرض خطر بسته شدن قرار می‌گیرد. از این رو، طی شش هفته پیش از برگزاری امتحانات هماهنگ خواندن و نوشتن انگلیسی، آموزگاران توسط مدیر مدرسه آموزش داده می‌شوند که تمام زمان کلاس را به تمرین برای آمادگی در امتحانات اختصاص دهند و دیگر موضوعات را به کلی نادیده بگیرند. (یک استراتژی که با حسن تعبیر می‌توان آن را «تراز برنامه درسی» نامید.) «آموزش به آزمون» همچون شوخی آماری، راهی است که در آن نهادها منحرف می‌شوند؛ تلاشی که در آن نهادها از هدف اصلی خود (آموزش) منحرف می‌شود و به دنبال اهداف معیار هستند به این دلیل که بقای‌شان وابسته به این اهداف است.

اثرات منحرف‌کننده عملکرد معیارها دست‌کم به همان اندازه در سراسر آتلانتیک نیز احساس می‌شود. [۱] آنجا هم یک سریال تلویزیونی دیگر توسط یک متخصص سابق درباره پدیده‌ای مشابه و بر اساس واقعیت ساخته شد. سریال بدن‌ها<sup>۱</sup> که توسط جد مرکوریو<sup>۲</sup>، پزشک سابق بیمارستان نوشته شده و در بخش زنان و زایمان بیمارستانی در یک کلان‌شهر در جریان است. در قسمت اول، یک جراح ارشد تازه‌وارد روی یکی از بیماران مبتلا به چند بیماری پیچیده عملی انجام می‌دهد که پس از این عمل، بیمار می‌میرد. پس از این رویداد رقیبش به او توصیه‌ای می‌کند: «یک جراح ارشد از هوش برتر خود استفاده می‌کند تا از هر شرایطی که ممکن است توانایی برتر او را در معرض آزمون قرار دهد، جلوگیری کند.» این یعنی که او از موارد دشوار همچون راهی برای حفظ میزان موفقیت خود پرهیز کند. استراتژی کلاسیک «سرشیر گرفتن»<sup>۳</sup> که به معنای پرهیز از موارد پرخطری است که می‌تواند تأثیر منفی روی میزان عملکرد داشته باشد. هزینه این تاکتیک این است که بیماران پرخطر برای جراحی، رها می‌شوند تا زیر جراحی نمیرند. «بدن‌ها» یک درام پزشکی است، اما پدیده‌ای را به تصویر می‌کشد

1. Bodies

2. Jed Mercurio

3. Creaming

که در جهان واقعی وجود دارد. مطالعات متعددی نشان داده‌اند که برخی از جراحان - برای مثال زمانی که بر اساس میزان موفقیت‌شان رتبه‌بندی می‌شوند یا به آنها پاداش داده می‌شود- از جراحی بیمارانی با وضعیت‌های پیچیده یا بحرانی حذر می‌کنند. به استثنای موارد دشوارتر - مواردی که احتمال بروز پیامدهای ضعیف‌تری را شامل می‌شود- سایر موارد میزان موفقیت جراحان را افزایش می‌دهد و در نتیجه موجب رشد و بهبود معیارها و اعتبارشان و میزان پاداش‌شان می‌شود. البته هزینه این موارد را بیماران محروم با از دست دادن زندگی خود می‌پردازند اما مرگ آنها در معیارها نشان داده نمی‌شود.

همان‌گونه که خواهیم دید بازی شاخص‌ها و معیارها در هر حوزه‌ای اتفاق می‌افتد: در اداره پلیس، در آموزش ابتدایی، راهنمایی و متوسطه، در پزشکی، در سازمان‌های غیرانتفاعی و البته در کسب‌وکار. بازی تنها بخشی از مشکلاتی است که به صورتی گریزناپذیر در نتیجه استفاده از معیارهای عملکرد به عنوان بنیانی برای پاداش یا تنبیه بروز می‌کند. چیزهایی هستند که می‌توان اندازه‌گیری‌شان کرد و چیزهایی هستند که ارزش اندازه‌گیری دارند اما آنچه اندازه‌گیری می‌شود همیشه ارزش آن را ندارد؛ آنچه اندازه‌گیری می‌شود ممکن است با آنچه ما به‌راستی به دنبال دانستن آن هستیم، ارتباطی نداشته باشد. هزینه‌های اندازه‌گیری ممکن است بیشتر از مزایای آن باشد، چیزهایی که اندازه‌گیری می‌شود ممکن است تلاش ما را از چیزهایی که برای ما به‌راستی اهمیت دارد، دور کند. می‌توان گفت اندازه‌گیری می‌تواند دانش تحریف‌شده به ما بدهد؛ دانشی که چه‌بسا قابل اطمینان به نظر برسد اما در حقیقت گمراه‌کننده است.

ما در عصر پاسخگویی اندازه‌گیری‌شده، پاداش برای عملکرد اندازه‌گیری‌شده و اعتقاد به شناساندن و ترویج این معیارها از طریق «شفافیت» زندگی می‌کنیم. اما شناسایی پاسخگویی با معیارها و با شفافیت گمراه‌کننده است. پاسخگویی باید به معنای مسئولیت‌پذیری در قبال اقدامات مشخص باشد اما با نوعی ترفند زبانی، پاسخگویی به

معنای نشان دادن موفقیت از طریق اندازه‌گیری استاندارد شده است؛ گویی آنچه اهمیت دارد به‌راستی قابل محاسبه و شمارش است. فرض دیگری که اغلب مورد تایید قرار می‌گیرد این است که «پاسخگویی» خواهان آن است که اندازه‌گیری عملکرد علنی شود و این به معنای «شفافیت» است. به نظر می‌رسد ثبات معیار، یک فشار غیرقابل مقاومت برای اندازه‌گیری عملکرد، عمومی‌سازی آن و پاداش برای آن است، هرچند در بیشتر موارد شواهدی وجود دارد که نشان می‌دهد این امر چندان درست هم انجام نمی‌شود.

خواهیم دید که اندازه‌گیری اگر به‌درستی استفاده شود می‌تواند مفید باشد. این گزاره درباره شفافیت نیز صدق می‌کند اما این موارد می‌تواند تحریف‌شده، همراه‌کننده، تغییر یافته، منحرف‌شده و دلسردکننده نیز باشد. در عین حال که ما مجبور به زندگی در عصر اندازه‌گیری هستیم، در حقیقت در عصر اندازه‌گیری نادرست، اندازه‌گیری بیش از حد، اندازه‌گیری همراه‌کننده و اندازه‌گیری غیرمولد زندگی می‌کنیم. این کتاب درباره بدی و زیان اندازه‌گیری نیست، بلکه درباره نتایج منفی ناخواسته تلاش برای جایگزین کردن اندازه‌گیری استاندارد عملکرد در راستای قضاوت افراد بر اساس تجربه است. مشکل، اندازه‌گیری و معیار نیست بلکه اندازه‌گیری بیش از حد و نامناسب و ثبات معیار است.

اغلب به ما گفته می‌شود که گردآوری معیارهای عملکرد اندازه‌گیری‌شده و سپس آنها را در دسترس عموم قرار دادن، روشی برای بهبود عملکرد نهادهای ماست. در هیچ‌کجا بیشتر از حوزه پزشکی درباره پاسخگویی، معیارهای عملکرد و شفافیت تبلیغ نشده است و بنابراین قابل درک است که در هیچ‌عرضه دیگری به اندازه این حوزه، خطر و زیان نیز بیشتر نباشد. حوزه بهداشت نه تنها ۱۷ درصد از اقتصاد ایالات متحده آمریکا را تشکیل می‌دهد بلکه همین حوزه است که زندگی افراد را در معرض خطر قرار می‌دهد. مطمئناً منطق حکم می‌کند که اندازه‌گیری عملکرد هم دلارهای بیشتری را ذخیره کند و هم زندگی‌های بیشتری را نجات دهد.

گردآوری اطلاعات استاندارد درباره میزان موفقیت جراحان یا میزان نجات بیماران پذیرش شده در بیمارستان‌های مشخص می‌تواند مفید باشد. اگر به پزشکان یا بیمارستان‌ها توسط سازمان‌های دولتی یا بیمه‌های خصوصی بر اساس میزان موفقیت‌شان در زنده نگه‌داشتن بیمار پاداش داده شود، چنین اندازه‌گیری می‌تواند موجب ایجاد انگیزه در مراقبت و درمان بهتر شود. اگر میزان موفقیت پزشکان و بیمارستان‌ها در دسترس عموم قرار بگیرد، در نتیجه شفافیت به مردم اجازه خواهد داد تا میان پزشکان و بیمارستان‌ها دست به انتخاب بزنند. به طور کلی معیارها، پاسخگویی و شفافیت درمانی برای دردها و آلام حرفه پزشکی ارائه می‌شود. اما چه اشتباهی می‌تواند انجام شود؟

یک معامله خوب؛ همان‌گونه که پیشتر دیده‌ایم زمانی که امتیازات جراحان به منزله معیاری برای پاداش یا تنبیه‌شان مورد استفاده و تحت چنین بررسی موشکافانه‌ای قرار می‌گیرد، آنها نیز همچون دیگران فعالیت‌ها را استاندارد خود را انجام می‌دهند و خود را درگیر کارهای پرخطر نمی‌کنند. زمانی که بیمارستان‌ها بر اساس درصد بیمارانی که ۳۰ روز پس از عمل جراحی از دنیا رفتند مجازات می‌شوند، گاهی بیماران به مدت ۳۱ روز زنده نگه داشته می‌شوند؛ از این رو، آمار مرگ‌ومیر در معیارهای بیمارستان منعکس نمی‌شود. [۲] در انگلستان در تلاشی برای کاهش زمان انتظار در بخش‌های اورژانس، دپارتمان سلامت سیاستی اتخاذ کرد که مطابق آن بیمارستان‌ها با زمان انتظاری بیش از چهار ساعت جریمه می‌شدند. این برنامه دست‌کم در ظاهر موفقیت‌آمیز بود اما در واقع برخی بیمارستان‌ها با نگه داشتن بیماران در صف آمبولانس‌ها آنها را پشت درهای بسته بیمارستان منتظر گذاشتند تا زمانی که کارکنان اطمینان یابند که بیماران تا چهار ساعت پس از ورود پذیرش خواهند شد. [۳]

ما به این مسائل در حوزه سلامت به طور عمیق‌تر نیز پرداخته‌ایم. اما نکته جالب توجه این است که چنین مشکلاتی در حوزه مراقبت‌های بهداشتی و سلامت در بسیاری از نهادهای دیگر نیز وجود دارد - در آموزش

تمام مقاطع تحصیلی و کالج، در اداره پلیس و دیگر خدمات عمومی، در کسب‌وکار و مالیه و در سازمان‌های خیریه. افرادی که در این حوزه‌ها مشغول کارند، چنین مشکلاتی را در نهادهای خود احساس می‌کنند و دانشمندان علوم اجتماعی این مشکلات را در یک یا چند حوزه دیگر نیز بررسی و تشریح کرده‌اند. آنچه تا اندازه زیادی مورد غفلت قرار گرفته، تکرار نتایج منفی ناخواسته مشابه معیارهای عملکرد، پاسخگویی و شفافیت در میان طیف گسترده‌ای از نهادهاست.<sup>[۴]</sup>

همچون بسیاری از بینش‌های دیگر، زمانی که از ثبات معیار<sup>۱</sup> آگاه شدید، این احتمال هست که آن را نه تنها در درام‌های تلویزیونی بلکه در جاهای دیگر بیابید.

کلمات کلیدی ثبات معیار در گوشه و کنار ما هستند. گوگل انگرام<sup>۲</sup> - که به سرعت در میان هزاران کتاب اسکن‌شده و دیگر نشریات جست‌وجو می‌کند- تصویری خشن اما پرمخاطب از تغییرات فرهنگ و جامعه به ما ارائه می‌دهد. پارامترها را بر مبنای سال تنظیم و یک اصطلاح یا عبارت را تایپ کنید، نموداری ظاهر می‌شود که تکرار کلمات را از ۱۸۰۰ تا امروز نشان می‌دهد. واژه «پاسخگویی» را تایپ کنید؛ خطی را خواهید دید که در حوالی ۱۹۶۵ شروع می‌کند به شکل منحنی رو به بالا رفتن و پس از ۱۹۸۵ با یک شیب فزاینده بالاتر می‌رود. درست مانند «معیارها»، که استفاده زیاد از آن در حدود سال ۱۹۸۵ آغاز می‌شود. جست‌وجو نشان می‌دهد که عبارت «محک‌زنی<sup>۳</sup>» نیز مانند «شاخص‌های عملکرد» از الگوی مشابهی پیروی می‌کند.

کتاب حاضر این استدلال را مطرح می‌کند که هر چند معیارها به طور بالقوه ابزارهای ارزشمندی هستند اما اغلب درباره حسن و درستی معیارهای

1. Metric fixation

2. Google's Ngram

این ابزار امکان جست‌وجو در میان پرتعدادترین کلمات به کار رفته در کتاب‌های گردآوری شده در کتابخانه مجازی گوگل را فراهم می‌کند. قابلیت دیگر این سرویس امکان جست‌وجو و تعداد تکرار هر عبارت در متن کتاب‌های لاتین در ۵۰۰ سال اخیر است. علاوه بر این، می‌توان در یک بازه زمانی مشخص، در رابطه با یک جمله یا عبارت خاص، در طول سخنرانی‌های افراد مطرح جهان نیز تحقیق کرد. نتایج تمام جست‌وجوها بر اساس یک نمودار ارائه خواهند شد.

3. Benchmarks

پاسخگویی زیاده‌روی و هزینه‌های آنها نیز دست‌کم گرفته شده است. این کتاب نیز یک سبب‌شناسی و تشخیص و در عین حال یک پیش‌آگاهی برای چگونگی پرهیز از ثبات معیارها و تسکین دردهای آن ارائه می‌دهد.

بارزترین ویژگی ثبات معیارها این است که خواست و اشتیاق همراه با قضاوت مبتنی بر تجربه با اندازه‌گیری استاندارد جایگزین می‌شود. زیرا قضاوت در حکم یک امر شخصی، ذهنی و خویش‌دوستی درک می‌شود. در عوض، معیارها قرار است اطلاعاتی قابل‌اطمینان و عینی ارائه دهد. استراتژی برای بهبود کارایی نهادی این است که به افراد دارای بالاترین معیارها یا افرادی که توانسته‌اند به رتبه‌بندی و اهداف برسند، پاداش داده شود و افرادی که عقب مانده‌اند تنبیه شوند. سیاست‌ها بر اساس این فروض برای چندین دهه رو به جلو حرکت کرده‌اند و همان‌گونه که شیب رو به رشد نمودارهای انگرام نشان می‌دهد، این واقعیت مفروض رو به جلو در حال حرکت است.

محض اطمینان باید گفت که در موقعیت‌های بسیاری تصمیم‌گیری طبق اندازه‌گیری استاندارد بر قضاوت مبتنی بر تجربه و تخصص شخصی ارجحیت دارد. تصمیمات مبتنی بر داده‌های بزرگ زمانی سودمند هستند که تجربه هر فرد متخصص به طور احتمالی در راستای بسط و گسترش یک حس شهودی یا اندازه قابل‌اطمینان از اثربخشی بیش از حد محدود شده باشد. برای مثال زمانی که یک پزشک با علائم یک اختلال نادر روبه‌رو می‌شود، بهتر است به او توصیه شود که بر معیارهای استاندارد حاصل از جمع‌آوری نمونه‌های فراوان اتکا کند. چک‌لیست‌ها - رویه‌های استاندارد برای نحوه اقدام در شرایط معمول - نشان داده‌اند که در حوزه‌های متنوعی مانند خطوط هوایی و پزشکی ارزشمند هستند. [۵] همان‌گونه که در کتاب مانیبال ذکر شده، تجزیه و تحلیل آماری گاهی به این کشف نائل می‌شود که ویژگی‌های قابل اندازه‌گیری واضح اما نادیده گرفته‌شده نسبت به درک شهودی شناخته‌شده

مبتنی بر تجارب اندوخته از اهمیت بیشتری برخوردار است. [۶]

بنابراین در صورت استفاده عاقلانه، اندازه‌گیری ویژگی‌هایی که پیش از این اندازه‌گیری نشده‌اند، می‌تواند مزایای مهمی در بر داشته باشد. تلاش برای

اندازه‌گیری عملکرد - همان‌گونه که خواهیم دید با مشکلاتی روبه‌روست - ذاتاً مطلوب است. اگر آنچه به طور واقعی اندازه‌گیری می‌شود، یک نماینده معقول برای آنچه برای اندازه‌گیری تعیین شده باشد و اگر این اندازه‌گیری همراه با قضاوت باشد، در این صورت می‌تواند به متخصصان و شاغلان برای ارزیابی عملکرد خود هم در سطح فردی و هم در سطح سازمانی کمک کند. اما مشکل از جایی آغاز می‌شود که این اندازه‌ها بدل به معیارهایی برای پاداش و تنبیه می‌شود؛ در واقع زمانی که معیارها به مبنای پرداخت مبتنی بر عملکرد یا رتبه‌بندی‌ها تبدیل شود.

طرح‌های عملکرد اندازه‌گیری‌شده به طور فریبنده‌ای جذابند زیرا در بیشتر مواقع خود را با کشف آشکارترین موارد اشتباه یا غفلت «اثبات» می‌کنند اما سپس این کشفیات به همه موارد تعمیم داده و ابزارهایی که برای کشف سوءرفتار واقعی مناسبند به ابزارهایی برای اندازه‌گیری کلیه عملکردها تبدیل می‌شود. یافته‌های اولیه اندازه‌گیری عملکرد ممکن است عملکردهای ضعیف را به سمت بهبود یا حذف از بازار سوق دهد اما در بسیاری از موارد، گسترش اندازه‌گیری استاندارد ممکن است ناشی از ابزارهای نقصان‌یافته یا حتی غیرمولد باشد؛ حرکت از راه‌حل‌های معقول تا جنون معیارها. اما مهم‌تر از همه این است که اندازه‌گیری هنگامی که بر آن است تا یک چیز غیرقابل اندازه‌گیری را اندازه‌گیری کند و یک چیز غیرقابل کمی را به صورت کمی درآورد، شاید نقشی غیرمولد داشته باشد.

به عبارتی منافع واقعی قدرت، پول و موقعیت در خطرند. ثبات معیارها موجب می‌شود که منابع از تولیدکنندگان خط مقدم به سوی مدیران، رئیس‌ان و افرادی که داده‌ها را جمع‌آوری و دستکاری می‌کنند، منحرف شود. زمانی که معیارها توسط مدیران بسان ابزاری برای کنترل متخصصان مورد استفاده قرار می‌گیرد، اغلب موجب بروز تنش میان مدیرانی می‌شود که به دنبال اندازه‌گیری عملکرد و پاداش بر مبنای آن هستند و اخلاق و شخصیت انسانی متخصصان (پزشکان، پرستاران، نیروهای پلیس، آموزگاران، استادان دانشگاه و غیره) را زیر سوال می‌برد. این در حالی است که اخلاق حرفه‌ای مبتنی است بر تسلط به مجموعه‌ای متشکل از دانش تخصصی کسب‌شده طی

یک فرآیند طولانی مدت در آموزش، استقلال و مسلط بودن بر کار، تطبیق با یک گروه حرفه‌ای و داشتن حس مسئولیت نسبت به همکاران، ارزش بالای پاداش‌های ذاتی و تعهد به منافع مشتریان فارغ از توجه به هزینه. [۷]

این تنش گاهی ضروری و خوشایند است زیرا اخلاق حرفه‌ای تمایل دارد تا مسائل مربوط به هزینه و هزینه-فرصت را نادیده بگیرد. به این معنا که متخصص تمایل دارد بدون در نظر گرفتن محدودیت منابع یا استفاده متناوب از منابع، فقط مزایای خدمات ارائه‌شده خود را ببیند. متخصصان علاقه‌ای به اندیشیدن درباره هزینه‌ها ندارند و افراد عادی اندازه‌گیری معیارها را انجام می‌دهند. هنگامی که این دو گروه (حرفه‌ای‌ها و افراد عادی) با هم کار می‌کنند، نتیجه می‌تواند برای هر دو رضایت‌بخش‌تر باشد زیرا هنگامی که این دو در برابر هم قرار می‌گیرند، نتیجه کشمکش و تضعیف روحیه است.

هرچند منافع مشخصی وجود دارد و حائز اهمیت است اما گاهی همین امر موجب می‌شود که معیارهای معقول به جنون معیاری تبدیل شود، دلیل این امر در پذیرش غیرقابل انتقاد ایدئولوژی مبتنی بر معیار نهفته است. مانند هر فرهنگی، فرهنگ پاسخگویی معیاری نیز دارای اصطلاحات مقدس انکارناپذیر و نقاط کور خاصی است. [۸] هرچند آنچه امروز حکمفرماست این است که عیوب تمایلی به جلب توجه ندارد.

ممکن است از اینکه یک تاریخ‌نگار چگونه کتابی درباره استبداد معیارها نوشته است، شگفت‌زده شوید. زمانی این اتفاق افتاد که من متوجه سیر تکاملی نگران‌کننده‌ای در تجربه حرفه‌ای خودم شدم که منعکس‌کننده الگوهای بزرگ‌تر در جامعه بود. نارضایتی‌های خرد منجر به تحلیل در سطح کلان می‌شود، زیرا من متوجه شدم که الگوهای فرهنگی که به حوزه محدود حرفه من آسیب رسانده‌اند، بسیاری از نهادهای معاصر را نیز منحرف کرده‌اند.

من به واسطه تجربه‌ام در جایگاه رئیس دپارتمان یک دانشگاه خصوصی جذب این موضوع شدم. چنین شغلی جوانب متعددی دارد: راهنمایی اعضای هیات علمی و کمک به آنها برای اینکه به منزله آموزگار یا

پژوهشگر رشد کنند، استخدام هیات علمی جدید، تلاش برای حصول اطمینان از آموزش دوره‌های لازم، حفظ روابط با روسا و دیگران در مدیریت دانشگاه. این مسئولیت‌ها مضاف بر وظایف من در نقش عضو هیات علمی بود: تدریس و پژوهش و پیگیری دائم حوزه‌های کاری‌ام. با وجود همه این وظایف، من کاملاً خرسند بودم. زمان اختصاص یافته به اندیشیدن درباره اعضای هیات علمی و کار با آنها موجب شد تا آنها آموزگاران و پژوهشگران بهتری شوند. من به گستره و کیفیت دوره‌هایی که تدریس می‌کردم افتخار می‌کنم و روابط با سایر بخش‌ها نیز خوب بود؛ در مجموع تدریس و پژوهش و نوشتن دشوار اما رضایت‌بخش بود. سپس چیزهایی شروع به تغییر کرد. مانند همه کالج‌ها و دانشگاه‌ها، موسسه ما نیز هر دهه توسط یک مجموعه معتبر یعنی کمیسیون آموزش عالی ایالت‌های مرکزی مورد ارزیابی قرار می‌گرفت. گزارشی منتشر شد و در این گزارش درخواست شده بود که معیارهای بیشتری در نظر گرفته شود و در آینده به منزله بنیانی برای «ارزیابی» مورد استفاده قرار گیرد - یک اصطلاح رایج مد روز در آموزش عالی که معمولاً به معنای ارزیابی بیشتر عملکرد است. خیلی زود متوجه شدم که برای اطلاعات آماری درباره فعالیت‌های دپارتمان زمان بسیار زیادی صرف پاسخ به پرسشنامه‌ها می‌کنم؛ زمانی که باید صرف پژوهش و تدریس و آموزش اعضای هیات علمی می‌شد. به این ترتیب مقیاس‌های جدیدی برای ارزیابی دستاوردهای رشته‌های دانشگاهی و فارغ‌التحصیلان ارائه شد - مقیاس‌هایی که هیچ بینش سودمندی به ابزار پیشین اندازه‌گیری ما یعنی نمرات نمی‌افزودند. من روی روشی برای انجام سریع این کار بدون گرفتن زمان زیاد از اعضای هیات علمی کار کردم و این امر به‌سادگی و با قرار دادن نمراتی که اعضای هیات علمی کسب کرده‌اند در چهار دسته مختلف و ایجاد مقیاسی با هدف ارزیابی انجام می‌شود. در طول زمان، گردآوری و پردازش اطلاعات، به نوبه خود، نیازمند این است که دانشگاه متخصصان داده بیشتری استخدام کند (این امر تا آنجا پیش رفته که یک معاون رئیس دانشگاه نیز به منظور ارزیابی این معیارها منصوب

شده است). بعضی گزارش‌های آنها حقیقتاً مفید بودند: برای مثال در تولید صفحات گسترده که میانگین نمرات کسب‌شده در هر دوره را نشان می‌دهد. اما بیشتر اطلاعات استفاده واقعی نداشتند و در حقیقت هیچ‌کس آنها را نمی‌خواند. با این حال هنگامی که فرهنگ مستندسازی عملکرد رواج یافت، روسای دپارتمان نیز خود را در یک رقابت تسلیحاتی داده‌ها یافتند. من دپارتمان را به سوی یک خودارزیابی سالانه دپارتمانی سوق دادم که همان‌گونه که مشخص شد، یک تمرین مفید به شمار رفت. اما پیش از ارسال آن به زنجیره بوروکراتیک از من خواسته شد تا ضمایم آماری بیشتری به آن اضافه کنم، زیرا اگر چنین نمی‌کردم گزارش دیگر دپارتمان‌ها نسبت به گزارش ما دقیق‌تر به نظر می‌رسید. یکی از اعضا -یک پژوهش‌گر ارشد قابل اعتماد- تقریباً تمام تابستان خود را وقف ایجاد پوشه‌ای پر از داده کرد و در نهایت آن را با نمودارهای رنگی در راستای تلاش برای اقناع رئیس دانشگاه که نیازمند این گزارش برای پر کردن شکاف دپارتمان خود است، تکمیل نمود.

تجربه من آزردهنده بود نه خردکننده، یک خراش بود و نه یک ضربه سهمگین اما همین خراش من را برانگیخت تا به طور عمیق‌تری درباره نیرووهایی که ما را به سوی این انحراف بیهوده زمان و تلاش سوق می‌دهند کنکاش کنم. کمیسیون ایالات مرکزی که انگیزه برای داده‌های بیشتر از آن نشأت گرفته است، زیر نظر وزارت آموزش ایالات متحده عمل می‌کند. این وزارتخانه به رهبری مارگارت اسپلینگز<sup>۱</sup>، کمیسیونی درباره آینده آموزش عالی تشکیل داده بود که گزارش خود را در سال ۲۰۰۶ با تأکید بر نیاز به پاسخگویی و گردآوری داده‌های بیشتر و هدایت سازمان‌های معتبر منطقه‌ای به سوی ایجاد «نتایج عملکرد» به عنوان هسته ارزیابی‌شان منتشر کرد.<sup>[۹]</sup> این سبک ارزیابی به نوبه خود، به آرامی ابتدا در کمیسیون ایالات مرکزی، سپس در دانشگاه ما و در نهایت در خود من رسوخ کرد. اسپلینگز رئیس شورای سیاست داخلی در زمان ریاست جمهوری جورج دبلیو بوش بود که قانون «هیچ کودکی نباید عقب بماند»<sup>۲</sup> را در سال ۲۰۰۱ به تصویب رساند. ابتدا فکر

1. Margaret Spellings

2. No Child Left Behind Act

می‌کردم که این قانون - که ارزیابی آموزگاران و مدارس را بر اساس نمرات دانش‌آموزان‌شان در امتحانات استاندارد گسترش می‌دهد - یک گام مثبت است اما با گذشت زمان، انتقادات تندی را از حمایت‌کنندگان سابق این قانون مانند معاون پیشین وزیر آموزش، دایان راویچ<sup>۱</sup>، شنیدم و آموزگاران کلاس که با من آشنایی داشتند به من گفتند هر چند عاشق تدریس هستند، اما تکثیر گروه‌بندی برنامه آموزشی به منظور حداکثر کردن عملکرد در امتحانات، شور و اشتیاق آنها را از بین برده است.

چنین حکایاتی من را به پژوهش بیشتر سوق داد که از ابزار فکری خود استفاده کنم و ریشه‌های فرهنگی و تاریخی گسترده‌تر و جلوه‌های معاصر فرهنگ عملکرد اندازه‌گیری‌شده و پاداش داده‌شده را که هر روز بیشتر در نهادهای ما رسوخ می‌کند، بررسی کنم. علایق حرفه‌ای من میان مرزهای تاریخ، اقتصاد، جامعه‌شناسی و سیاست بوده است. برای مدت زمانی طولانی به تاریخ علاقه‌مند بودم؛ آنچه «سیاست‌گذاری عمومی» نام دارد و کتابی نیز درباره آدام اسمیت با عنوان یک تحلیلگر سیاست‌گذاری عمومی<sup>۲</sup> منتشر کرده بودم. همچنین کتابی درباره رهیافت‌های محافظه‌کار نسبت به سیاست‌گذار عمومی نوشته بودم و مشخص شد بعضی اندیشمندان مانند مایکل اوکشات<sup>۳</sup> و فردریش هایک<sup>۴</sup> که درباره آنها نوشته بودم بینش‌های انتقادی نسبت به ستایش اغراق‌آمیز معاصر ما از عملکرد معیار ارائه داده‌اند. من علاقه‌مند به تاریخ سرمایه‌داری بوده‌ام، به‌ویژه راه‌هایی که در آن روشنفکران درباره پیش‌نیازهای اجتماعی، اخلاقی و سیاسی و نتایج کسب‌وکار تفکر می‌کنند. یک نگرانی مکرر میان روشنفکران مدرن غربی نسبت به افرادی که من درباره آنها نوشته بودم عبارت بود از سرریز بالقوه مهلک مفاهیم و گزاره‌ها از حوزه تجارت و علم اقتصاد به دیگر قلمروهای زندگی. همین‌طور تجربه شخصی من از نارضایتی حرفه‌ای موجب تحقق فرصتی شد و من را برانگیخت تا پژوهش‌های بیشتری انجام دهم و طیف گسترده‌ای از علایقم را ترسیم

1. Diane Ravitch

2. A Public Policy Analyst

3. Michael Oakeshott

4. Friedrich Hayek

کنم. روح حاکم بر این کتاب متعلق به مَتیو آرنولد<sup>۱</sup> و رابرت ک. مِرتون<sup>۲</sup> است؛ مَتیو که منتقد فرهنگ و ویکتوریاست و رابرت که به من آموزش داد تا به دنبال نتایج غیرمنتظره و ناخواسته کنش اجتماعی باشم و همچنین فرصت بورس تحصیلی را برای من ایجاد کرد. [۱۰]

هنگامی که من پژوهش درباره این موضوعات را آغاز کردم، کتاب یک جامعه‌شناس به نام راکش خورانا<sup>۳</sup> در مدرسه کسب‌وکار هاروارد با عنوان «از اهداف عالی تا دستان اجیر شده: تحول اجتماعی مدارس کسب‌وکار آمریکایی و وعده تحقق نیافته مدیریت به عنوان یک حرفه»<sup>۴</sup> چشمانم را بر تاریخ روشنفکری مدارس کسب‌وکار و تاثیر گسترده آنچه در این مدارس تدریس می‌شد باز کرد. این بینش‌ها من را به پژوهش‌های گسترده‌تر درباره تغییر فرهنگ و ایدئولوژی‌ها در حوزه مدیریت سوق داد و گاهی به سوی ماهیت نامطمئن این حوزه می‌برد که به خوبی در عنوان کتابی از آدرین وولدریج<sup>۵</sup> به نام «پزشکان جادوگر» به کار گرفته شده است (ویرایش دوم عنوان مهربانانه‌تری دارد: اربابان مدیریت).

من در طیف گسترده‌ای از ادبیات و تحقیقات علمی از اقتصاد و سیاست تا تاریخ، مردم‌شناسی، روان‌شناسی، جامعه‌شناسی، اداره امور عمومی و رفتار سازمانی جست‌وجو کردم. من از مطالعات علوم اجتماعی رفتار واقعی آموزگاران، استادان دانشگاه، پزشکان و پلیس‌ها در جهان واقعی استفاده زیادی کردم. در بررسی موضوع بورس تحصیلی رشته‌های مختلف، من از میزانی تمایل رشته‌های دانشگاهی به جدایی از یکدیگر و شکاف میان پژوهش دانشگاهی و عمل دنیای واقع تاثیر پذیرفتم. برای مثال متوجه این نکته مهم شدم که ادبیات اقتصادی اخیر درباره انگیزه‌ها و محرک‌ها تا چه اندازه، به رسمیت شناختن همان چیزهایی است که روان‌شناسان پیشتر کشف کرده‌اند. اما بیشتر چیزهایی هم که روان‌شناسان کشف کرده‌اند، خیلی پیش‌تر توسط

1. Matthew Arnold

2. Robert K. Merton

3. Rakesh Khurana

4. From Higher Aims to Hired Hands: The Social Transformation of American Business Schools and the Unfulfilled Promise of Management as a Profession

5. Adrian Wooldridge

مدیران با قضاوت کشف شده بود. اگرچه هنوز مجموعه بزرگی از پژوهش‌ها در حوزه‌های روان‌شناسی و اقتصاد، انگاره‌ها و اثربخشی توجه به عملکرد اندازه‌گیری شده را زیر سوال می‌برند، اما به نظر می‌رسد که ادبیات نسبت به توقف گسترش ثبات معیار اقدامات اندکی انجام داده است. [۱۱]

به همین دلیل من این کتاب را نوشتم. اندکی از آنچه این کتاب گفته کاملاً جدید است زیرا بیشتر در نتیجه در هم آمیختن پژوهش‌ها و بینش‌های نویسندگان دیگر است. بسیاری از سوءعملکردهای مرتبط با آنچه من آن را به اصطلاح «ثبات معیار» می‌نامم توسط پژوهش‌گران این حوزه یا حوزه‌های دیگر همچون آموزش، پزشکی، پلیس، واحدهای تجاری انتفاعی و غیرانتفاعی ثبت و تحلیل شده است. تعداد اندکی از دانشجویان رفتار سازمانی که در حوزه‌های نسبتاً تخصصی می‌نویسند، برخی الگوهای گسترده موفقیت و سوءعملکرد را تحلیل کرده‌اند. در حقیقت آنچه هیچ‌کس تاکنون انجام نداده است قرار دادن همه اینها کنار هم و در دسترس قرار دادن آن برای همه ماست که در این موسسات پیشرو کار می‌کنیم - از سیاستمداران گرفته که درباره سرنوشت نظام‌های آموزشی و پزشکی تصمیم می‌گیرند، تا اعضای هیات‌مدیره شرکت‌ها، متولیان دانشگاه‌ها و سازمان‌های غیرانتفاعی و همین‌طور افراد رده‌پایین (مانند روسای دپارتمان‌ها). به طور گسترده‌تر، این کتاب برای تمام افرادی است که به دنبال دانستن یکی از بزرگ‌ترین دلایل این هستند که چرا امروز بسیاری از سازمان‌های معاصر عملکرد خوبی و درخوری ندارند و با کاهش تولید، کارکنان‌شان را از خود ناامید می‌کنند.

با وجود اینکه محور این استدلال با باور همگانی درباره نهادهای معاصر زاویه دارد، هدف من این نبوده است که به شکل کاملاً نوین به سراغ آن بروم، بلکه می‌خواستم با یک خرد پالوده<sup>۱</sup> به آن پردازم. خوانندگانی که مشتاق طبقه‌بندی استدلال در برخی چارچوب‌های ایدئولوژیک موجودند، ناامید خواهند شد زیرا این استدلال نه تنها از رشته‌های مختلف که از گرایش‌های سیاسی متعددی نیز طرح شده است. من شواهد و بینش‌ها را از هر کجا که یافت شده‌اند ترسیم کرده‌ام. در نهایت امیدوارم خوانندگان با ذهنی باز به کتاب نزدیک شوند.

## سپاسگزاری

نویسنده

این کتاب برگرفته از تجربه‌هایی است که در مقدمه به آنها پرداخته شد. من ابتدا نظرات خود را در این باره در مقاله‌ای با عنوان «هزینه‌های پاسخگویی» که در مجله امریکن اینترست<sup>۱</sup> (سپتامبر-اکتبر ۲۰۱۵) به چاپ رسید بیان کردم. از آدام گارفینکل ویراستار مجله ممنونم که مشتاقانه نوشته‌ام را پذیرفت، با مهارت آن را ویرایش کرد و مجوز انتشار این کتاب را گرفت.

از الیوت ا. کوهن، رافائل کوهن، هارولد جیمز، ناتان لویتان، ایسه پارکر، توماس پترسون، آویل توکر و آدرین وولدریج سپاسگزارم که با تشویق‌ها و اندرزهای‌شان موجب شدند مقاله‌ام را به چنین کتابی تبدیل کنم. جوئل برنر، آرنولد کلینگ و مرحوم کریستوفر کوبراک آنقدر مهربان بودند که متن را بخوانند و پیشنهادهای خوبی برای بهتر شدنش دهند. همچنین از همکارانم کارولاین شرمن و استفان وست در دانشگاه کاتولینگ بسیار آموختم. از همه آنها و دوستان و همکاران دیگری که به من پیشنهاد منابع و چارچوب پژوهش را دادند، بسیار سپاسگزارم. همچنین قدردانی می‌کنم از چهار منتقد ادبی ناشناس در انتشارات دانشگاه پرینستون که تنها انگیزه درونی و ذاتی داشتند و زمان خود را

1. The American Interest

سخاوتمندانه به من دادند و به من کمک کردند تا استدلال‌های خود را بازبینی و اصلاح کنم.

بخش‌هایی از این کتاب در سمینارهایی ارائه شده است که توسط راجشری آگاروال و دیوید سیسیلیا در مدرسه کسب‌وکار اسمیت دانشگاه مریلند یا در کالج پارک و توسط دانیل کلین در دانشکده اقتصاد دانشگاه جرج میسون یا توسط کترین جنسن در کنفرانس اعضای هیات علمی دانشکده تاریخ دانشگاه کاتولیک آمریکا برگزار شده است. من از بازخورد هر یک از این مناسبت‌ها بسیار سود برده‌ام.

چندین سال است که از مزایای یک سازمان غیررسمی اما دارای سمینار و رفتار سازمانی بهره برده‌ام؛ متشکل از همسر، سه فرزندمان و همسران‌شان. آنها بینش‌های خود را در رابطه با کارکردها و اشکال‌های سازمان‌هایی که در آنها کار می‌کنند شامل دولت، آموزش و پرورش و پزشکی با من به اشتراک گذاشتند. به همه آنها مدیونم، هرچند هیچ کدام مسئول استفاده‌های من از مشاهدات‌شان نیستند. الی مولر توجه مرا به مضامین بزرگ‌تر سریال شنود جلب کرد (قسمت‌هایی که می‌توان یک فصل از آنها نقل قول آورد و درباره‌شان شعر گفت)، و تجزیه و تحلیل دقیق او درباره پویایی‌های نهادی در صفحات این کتاب بازتاب یافته است، همچنین در ویرایش کتاب نیز بسیار کمک کرد. از دکتر جوزف مولر تشکر می‌کنم که راهنمای من در ادبیات پزشکی و سلامت و همچنین توصیه‌های هوشمندانه‌اش در لحن و جهت‌گیری کتاب بسیار ارزشمند بود. همسرم شارون، نخستین خواننده و ویراستار هر فصل بود و بسیاری از ایده‌های کتاب در مکالمات روزانه ما متولد یا اصلاح شدند (ما درباره لذت بی‌حدو حصر مادر بزرگ و پدر بزرگ بودن صحبت نمی‌کردیم - هرچند این موضوع کتاب دیگر است).

کامل کردن این کتاب با حمایت‌های مادر و پدرم ساده بود. بسیار از گفتن این موضوع غمگینم که این آخرین پروژه‌ای بود که توانستم درباره‌اش با پدرم، هنری مولر صحبت کنم که زمان تکمیل این کتاب از دنیا رفت، روحش قرین آرامش. مادرم بلا مولر همچنان یک منبع

قدرتمند از هوشمندی، دلگرمی و شوخ‌طبعی در زندگی‌ام است. از جسیکا یائو از انتشارات دانشگاه پرینستون برای هدایت پروژه و تبدیل آن از نسخه خطی به کتاب و لیندا ترویلو برای نسخه برداری از نسخه خطی بسیار سپاسگزارم.

در ربع قرن گذشته، پیترو جی. دافری نه تنها ویراستار من که همفکر و همراه من بوده و مرا با جریان مداومی از کتاب‌ها و ایده‌ها تغذیه کرده و به نوشتن کتاب برای مخاطبان در خارج و داخل دانشگاه تشویق کرده است. او مرا متقاعد کرد که ایده‌هایم را ابتدا به صورت مقاله و سپس در کتاب‌هایم به صورت گام به گام منتشر کنم. این کتاب به او تقدیم می‌شود؛ با سپاس و مهر بسیار.

#### مترجم

از آقای حسین سلاح‌ورزی، نائب رئیس اتاق بازرگانی، صنایع، معادن و کشاورزی که با روی باز پیشنهاد ترجمه این کتاب را پذیرفتند و زمینه انجام این کار را فراهم کردند، سپاسگزارم.

بخش اول

# استدلال



## فصل ۱

## استدلال در چند کلمه

یک الگوی فرهنگی در دهه‌های اخیر فراگیر شده است و طیف در حال گسترشی از نهادها را در بر می‌گیرد. بسته به سلیقه می‌توان آن را یک «الگوی رفتاری»، «معرفت‌شناسی»، «گفتمان»، «پارادایم»، «نظام گفتاری خودتقویت‌کننده» [۱] یا به‌سادگی یک مُد فرهنگی نامید. این الگو واژگان و اصطلاحات ویژه خود را به همراه می‌آورد و بر روشی که مردم درباره جهان صحبت می‌کنند، یا درباره آن می‌اندیشند و بر اساس آن عمل می‌کنند، اثرگذار است. [۲] برای سادگی، آن را ثبات معیار می‌نامیم.

فرض کلیدی ثبات معیار درباره ارتباط میان اندازه‌گیری و پیشرفت است. گفته‌ای است که به اشتباه آن را به پزشک قرن نوزدهم لرد کلوین نسبت داده‌اند: «اگر نمی‌توانی آن را اندازه بگیری، نمی‌توانی آن را بهبود ببخشی.» در سال ۱۹۸۶ یک آموزگار مدیریت آمریکایی به نام تام پیترز<sup>۲</sup> از شعار «آنچه اندازه گرفته شود، انجام می‌شود» استقبال کرد که سنگ بنای باور به معیارهاست. [۳] با گذشت زمان برخی به این

1. Lord Kelvin

2. Tom Peters

نتیجه رسیدند که «هر چیزی که قابل اندازه‌گیری باشد، قابل پیشرفت است.» [۴]

هنگامی که طرفداران معیارها از «پاسخگویی» پشتیبانی می‌کنند، به طور ضمنی دو معنای واژه را در هم می‌آمیزند؛ از طرفی پاسخگو بودن به معنای «مسئول بودن» و از سوی دیگر به معنای «قابل‌شمارش بودن» است. حامیان «پاسخگویی» به طور معمول تصور می‌کنند که فقط با شمارش است که نهادها می‌توانند حقیقتاً مسئول باشند. از این رو عملکرد برابر است با آن چیزی که بتوان آن را به اندازه‌های استاندارد تقلیل داد. هنگامی که طرفداران معیارها، تقاضای «شفافیت» دارند اغلب به این نکته اشاره می‌کنند که درستی و راستی نیازمند تولید اطلاعات صریح و قابل‌مشاهده تا حد ممکن است. در نتیجه این امر تقاضا برای مستندسازی، شرح ماموریت و «تعیین هدف» بیشتر و بیشتر شد. [۵]

اجزای کلیدی ثبات معیار عبارتند از:

- باور به اینکه این احتمال وجود دارد و حتی مطلوب است که قضاوت بر اساس تجربه و استعداد شخصی جایگزین شاخص‌های متعدد مقایسه‌ای عملکرد بر اساس داده‌های (معیارهای) استاندارد شود؛

- باور به اینکه عمومی (شفاف) کردن این معیارها این اطمینان را می‌دهد که نهادها واقعا اهداف خود را انجام می‌دهند (پاسخگویی)؛

- باور به اینکه بهترین راه تشویق افراد در این سازمان‌ها قرار دادن پاداش و تنبیه برای عملکرد اندازه‌گیری‌شده آنهاست؛ پاداش‌هایی که یا پولی (پرداخت برای عملکرد) یا اعتباری (رتبه‌بندی) است. ثبات معیار، استمرار این باورهاست، به‌رغم اینکه زمانی که به کار گرفته می‌شوند دارای نتایج منفی ناخواسته هستند. [۶] این نتایج به این دلیل پیش می‌آیند که همه چیزهای مهم قابل اندازه‌گیری نیستند و بسیاری از چیزهای قابل اندازه‌گیری نیز اهمیت ندارند - یا بنا بر یک گفته آشنا: «نه همه چیزهایی که می‌توان شمرد، شمردنی هستند و نه

همه چیزهای شمردنی را می‌توان شمرد.» [۷]. بیشتر سازمان‌ها اهداف چندگانه‌ای دارند و آن هدفی که اندازه‌گیری می‌شود و پاداش می‌یابد، تمایل دارد تا حتی به قیمت قربانی کردن دیگر اهداف مهم سازمان در کانون توجهات باشد. به طور مشابه بسیاری از مشاغل نیز ابعاد چندگانه‌ای دارد و اندازه‌گیری ابعاد اندکی از آنها موجب ایجاد انگیزه‌هایی برای نادیده انگاشتن سایر ابعاد می‌شود. [۸] زمانی که سازمان‌ها متعهد به معیارها هستند، نسبت به این حقیقت نیز آگاه‌اند و به طور معمول اندازه‌گیری عملکرد بیشتری را به معیارهای پیشین اضافه می‌کنند - که این امر موجب شکل‌گیری مجموعه‌ای از داده‌ها می‌شود، داده‌هایی که فایده‌چندانی ندارند، در حالی که گردآوری آنها زمان و منابع بسیاری را به خود اختصاص داده است.

در این فرآیند، ماهیت کار به روش‌هایی تغییر می‌یابد که اغلب زیان‌بار است. افراد شاغل تمایل دارند تا نسبت به اهداف تحمیلی که با اخلاق حرفه‌ای و قضاوت‌شان مغایر است، اظهار انزجار کنند و به این ترتیب روحیه تضعیف می‌شود. بسیاری از افرادی که در دستکاری شاخص‌های عملکرد به روش‌های مختلف مهارت دارند، به شکلی گریزناپذیر برای سازمان‌های‌شان کاملاً ناکارآمد هستند. این افراد داده‌ها را سرهم‌بندی می‌کنند یا تنها با مواردی سروکار دارند که شاخص‌های عملکرد را بهبود می‌بخشد. آنها نمونه‌های منفی را گزارش نمی‌دهند و در موارد استثنایی شواهد را نیز جعل می‌کنند.

یک ویژگی تکراری ثبات معیار، پرداخت مبتنی بر عملکرد است که به افراد و سازمان‌ها پیشنهاد مشوق‌های مالی می‌دهد تا معیارهای قابل سنجش را شناسایی کنند. این کار ممکن است در سازمان‌هایی که تنها هدف‌شان سودآوری است جواب دهد، هرچند همان‌گونه که خواهیم دید حتی در این موارد نیز به ندرت اثرگذار است. این امر در سازمان‌هایی همچون مدارس، دانشگاه‌ها، اقدامات پزشکی و بیمارستان‌ها که افراد شاغل در آنها به سمت یک مأموریت آرمان‌گرایانه‌تر سوق داده می‌شوند به خوبی عمل نمی‌کند. بنابراین تا زمانی که پاداش به

عملکرد اندازه‌گیری شده گره خورده است، ثبات معیار، افراد را به بازی کردن دعوت می‌کند.

از آنجا که نظریه انگیزه و تشویق در راستای پرداخت مبتنی بر عملکرد اندازه‌گیری شده کمرنگ می‌شود، نتایج اغلب با انتظارات مغایرت دارد. الگوی معمول ناکارآمدی در سال ۱۹۷۵ توسط دو دانشمند علوم اجتماعی که در دو سوی متقابل اقیانوس اطلس مشغول به کار بودند کشف شد و به نظر می‌رسد هر یک در کشفی مستقل آن را ترسیم کرده‌اند. آنچه که «قانون کمپبل<sup>۱</sup>» نامیده شد، متعلق به روان‌شناس اجتماعی آمریکایی دونالد تی. کمپبل<sup>۲</sup> بود که باور داشت: «هر چقدر از شاخص اجتماعی کمی برای تصمیم‌گیری اجتماعی بیشتر استفاده شود، موضوع سنگینی فساد نیز بیشتر می‌شود و شرایط برای تحریف و فساد فرآیندهای اجتماعی که باید تحت نظارت قرار بگیرند نیز مناسب‌تر خواهد شد.» [۹] همین گزاره به شکلی دیگر توسط اقتصاددان انگلیسی با عنوان قانون گودارت<sup>۳</sup> ارائه شده است: «هر معیاری که برای کنترل (نظارت) استفاده شود غیرقابل اطمینان است.» [۱۰] به عبارت دیگر، هر چیزی که بتواند اندازه‌گیری و پاداش داده شود با آن بازی خواهد شد.

تلاش برای وادار کردن افراد به اینکه کار خود را با اهداف عددی از پیش تعیین‌شده وفق دهند، نوآوری و خلاقیت را - که ویژگی‌های ارزشمندی در بیشتر زمینه‌هاست - سرکوب می‌کند. این امر در بیشتر مواقع به طور گریزناپذیری به ارزیابی اهداف کوتاه‌مدت می‌انجامد تا بلندمدت.

در شرایطی که هیچ راه حل امکان‌پذیر واقعی برای یک مشکل وجود ندارد، گردآوری و انتشار داده‌های عملکرد به منزله ژستی از فضیلت کاربرد دارد. هیچ پیشرفت واقعی برای نشان دادن وجود ندارد اما تلاش در گردآوری و عمومی‌سازی داده‌ها افراد را از داشتن یک احساس حسن نیت اخلاقی، راضی و خشنود می‌سازد. به جای پیشرفت واقعی،

1. Campbell's Law

2. Donald T. Campbell

3. Goodhart's Law

پیشرفت اندازه‌گیری به چهره خیالی موفقیت بدل می‌شود. به «شکاف موفقیت»<sup>۱</sup> تحصیلی که مرتبط با این بحث است در بخش‌های بعدی کتاب می‌پردازیم.

از آنجا که به نظر می‌رسد شواهد عالی مبتنی بر کارایی این امر در بیشتر موارد درست نیستند، می‌توان گفت ثبات معیار دارای عناصر یک مکتب فکری است. مطالعاتی که نشان‌دهنده عدم اثربخشی آن است یا نادیده گرفته یا با این ادعا روبه‌رو می‌شوند که نیاز به داده‌های بیشتر و اندازه‌گیری بهتر وجود دارد. ثبات معیار تمایل به تقلید از علم دارد اما شبیه ایمان است.

همه این موارد قصد این ادعا را ندارد که بگویند اندازه‌گیری کاملاً بی‌فایده یا ذاتاً زیان‌آور است. یکی از اهداف این کتاب مشخص کردن زمان برای مفید بودن عملکرد معیارها - چگونگی استفاده از معیارها بدون ویژگی ناکارآمدی ثبات معیار - است.

فصل بعد، با عنوان «اشکالات تکراری» به طبقه‌بندی پرتکرارترین اشکالات در استفاده از معیارهای عملکرد می‌پردازد. تعریف و طبقه‌بندی این اشکالات موجب می‌شود تا بعدها ساده‌تر بتوانیم به آنها رجوع کنیم. سپس در بخش دوم نگاهی به ریشه‌های ثبات معیار خواهیم داشت و دلیل گسترش و ماندگاری آن را با وجود شکست‌های مکرر بررسی خواهیم کرد، همچنین در برخی منابع فلسفی عمیق‌تر، کاستی‌های آن را کاوش می‌کنیم. فصل سوم شامل مطالعات موردی اخیراً انجام‌شده و دربردارنده ثبات معیار است و به بررسی موفقیت‌ها و کاستی‌های آن در رشته‌های مختلف از جمله آموزش مقاطع تحصیلی ابتدایی، راهنمایی و متوسطه، پزشکی، پلیس، ارتش، تجارت، کمک‌های بشردوستانه و خارجی می‌پردازد. این مطالعات موردی بیشتر پیشنهادی هستند تا قطعی؛ به این معنا که با هر روشی که ثبات معیار توسط آن در هر حوزه خود را بروز می‌دهد، مواجه نمی‌شوند و در عوض نمونه‌های

1. Achievement gap

شکاف موفقیت به تفاوت عملکرد دانش‌آموزان و دانشجویان اطلاق می‌شود. این شکاف از طریق مقایسه نمرات، دوره‌های انتخابی و رتبه‌بندی‌های مختلف به دست می‌آید و در نمودارها و جداول نمایش داده می‌شود.

مشخصی از اشکالات مکرر و نتایج ناخواسته و همچنین نمونه‌های استفاده موفقیت‌آمیز از معیارها از درس‌آموخته‌های ما و به کارگیری آنها در دیگر حوزه‌ها را ارائه می‌دهند. در ادامه این بخش، بحث مختصری با موضوع شفافیت در حکم دشمن عملکرد در برخی حوزه‌های مشخص آمده است. در پایان فصل چهارم، تحلیل پیشین نتایج منفی ناخواسته ثبات معیار و چند رهنمود برای زمان و چگونگی استفاده از معیارها بدون تسلیم شدن در برابر ثبات معیار ارائه شده است.

## فصل ۲

## اشکالات تکراری

انگیزه ایجاد معیارها اغلب برخاسته از بهترین اهداف و به عنوان راه حل ارائه شده برای مشکلات واقعی است و در برخی موارد -همان گونه که خواهیم دید- معیارها وعده خود را در ارائه چنین راه حل‌هایی یا دست کم مشارکت در حل مشکلات محقق می‌کنند. اما پس از دهه‌ها تجربه در ارتباط با اثرات منفی معیارها، به دلیل اینکه ناکارآمدی معیارها با نفوذ به دیگر نهادها، آنها را تهدید می‌کند، باید بتوانیم مشکلات تکراری را پیش‌بینی کنیم. در اینجا فهرستی ارائه می‌شود که به شناسایی و یادآوری اشکالات کمک می‌کند. هر چند ممکن است ما این اشکالات را با هدف تحلیل مشخص و متمایز کنیم اما در جهان واقع با یکدیگر همپوشانی دارند. بیایید با مشکل تحریف اطلاعات آغاز کنیم:

**اندازه‌گیری ساده‌ترین چیزهای قابل اندازه‌گیری:** بشر یک تمایل ذاتی برای ساده‌سازی مشکلات با تمرکز بر ساده‌ترین عناصر قابل اندازه‌گیری دارد. [۱] اما آنچه به ساده‌ترین شکل اندازه‌گیری می‌شود، به ندرت مهم‌ترین چیز و در حقیقت، گاهی بی‌اهمیت است؛ این مورد اولین منبع ناکارآمدی معیار به شمار می‌رود.

یک موضوع کاملاً مرتبط عبارت است از اندازه‌گیری یک چیز ساده، وقتی خروجی مطلوب پیچیده است: بیشتر مشاغل و سازمان‌ها مسئولیت‌ها و اهداف چندگانه دارند. تمرکز اندازه‌گیری بیشتر روی یک مسئولیت یا هدف است که اغلب منجر به نتایج گمراه‌کننده می‌شود.

**اندازه‌گیری ورودی‌ها به جای خروجی‌ها:** اغلب ساده‌تر است که میزان هزینه‌ها و منابع تزیق شده به یک پروژه اندازه‌گیری شود تا نتایج تلاش‌ها. از این رو سازمان‌ها بیشتر آنچه را هزینه می‌کنند اندازه می‌گیرند تا آنچه را که تولید می‌کنند؛ به عبارتی بیشتر فرآیند را اندازه‌گیری می‌کنند تا تولید را.

**تخریب کیفیت اطلاعات از طریق استانداردسازی:** کمیت اغواکننده است، زیرا دانش را طبقه‌بندی و ساده‌سازی می‌کند. کمیت اطلاعاتی عددی ارائه می‌دهد که امکان مقایسه ساده میان افراد و نهادها را فراهم می‌کند. [۲] اما این ساده‌سازی چه‌بسا به تحریف بینجامد، زیرا مقایسه چیزها اغلب به این معناست که آنها از زمینه، تاریخ و معنای خود دور شده‌اند. [۳] نتیجه این است که اطلاعات، قطعی‌تر و معتبرتر از واقعیت به نظر می‌رسند: احتیاط‌ها، ابهامات و تردیدها از بین می‌روند و هیچ چیزی بیش از نمود دانش قطعی نسبت به بروز آن به شکل عددی انجام نمی‌شود. [۴]

قانون کمبل و قانون گودارت، خطارهایی درباره تلاش‌های گریزناپذیر بازی معیار می‌دهند، به‌ویژه زمانی که بیشتر در معرض خطر هستند. بازی معیارها شکل‌های مختلفی به خود می‌گیرد.

**بازی از طریق سرشیر گرفتن:** این مورد زمانی اتفاق می‌افتد که شاغلان اهداف ساده‌تری پیدا کنند یا مشتریانی با شرایط کم‌چالش‌تر را ترجیح دهند، این امر رسیدن به هدف معیار را ساده‌تر می‌کند اما مواردی استثنایی نیز هستند که موفقیت در آنها دشوارتر حاصل می‌شود.

#### 1. Creaming

اصطلاح سرشیر یا خامه گرفتن به موقعیتی اشاره دارد که کار یا فرآیند به صورت عمیق و دقیق انجام نمی‌شود و همه چیز در سطح باقی می‌ماند. (سطحی‌نگری). م

**بهبود اعداد با پایین آوردن استانداردها:** یک روش بهبود نمرات معیار، پایین آوردن معیار برای کسب نمره است. از این رو، مثلاً می‌توان با پایین آوردن استانداردهای قبولی، نرخ فارغ‌التحصیلی دبیرستان‌ها و کالج‌ها را افزایش داد یا خطوط هوایی با افزایش برنامه‌ریزی زمان پرواز پروازهای خود، عملکرد به موقع خود را بهبود می‌بخشند.

**بهبود اعداد از طریق حذف یا تحریف داده‌ها:** این استراتژی شامل نادیده گرفتن نمونه‌های دشوار یا دسته‌بندی نمونه‌ها به روشی است که آنها را از معیارها حذف کند. نیروهای پلیس می‌توانند با ثبت مجرمان به عنوان متخلفان یا با تصمیمی مبنی بر ثبت نکردن جرایم، آمار آن را «کاهش» دهند.

**تقلب کردن:** یک قدم فراتر از بازی معیارها، تقلب است؛ پدیده‌ای که افزایش وقوع تکرار آن ارتباط مستقیمی با خطرات معیار مورد بحث دارد. همان‌گونه که خواهیم دید، همان‌طور که قانون «هیچ کودکی نباید عقب بماند»، ضوابط جدید پیش روی مدارس برای نمرات آزمون دانش‌آموزان را مطرح کرد، آموزگاران و مسئولان در شهرهای مختلف با تغییر پاسخ‌های دانش‌آموزان به آزمون به این قانون واکنش نشان دادند.



بخش دوم

پیشینه



## فصل ۳

## خاستگاه‌های اندازه‌گیری و پرداخت مبتنی بر عملکرد

«پاسخگویی»، «معیارها» و «شاخص‌های عملکرد» به الگوهای رفتاری فرهنگی بدل شده‌اند. پذیرش این الگوها، نوید یک صندلی در قطار پیشرفت تاریخی را می‌دهد و هیچ سیاستمدار، رئیس اداره، دانشگاه یا مدرسه‌ای نمی‌خواهد از این قطار جا بماند. هنگامی که معیارها به پول رایج یک کشور بدل می‌شوند، استفاده نکردن از آنها خطر ورشکستگی را به همراه خواهد داشت. در نتیجه این امر، فشاری از جانب مقامات انتخابی و مدیران بنیاد برای پرداخت وجود دارد. چگونه و چرا این استبداد معیارها اتفاق می‌افتد؟

**برخی خاستگاه‌های پرداخت مبتنی بر عملکرد اندازه‌گیری شده**  
به نظر می‌رسد این ایده که «اگر بر اساس عملکرد اندازه‌گیری شده به سازمان‌های خارج از بازار آزاد پرداخت می‌شد، کارآمدتر بودند»، اولین بار توسط سیاست‌گذاران بریتانیا در دوره ویکتوریا مطرح شده

است. در سال ۱۸۶۲ یک عضو لیبرال پارلمان به نام رابرت لو<sup>۱</sup> که بر کمیته آموزش نظارت داشت، روش جدیدی برای تامین بودجه دولتی مدارس ارائه داد که مبتنی بر «پرداخت بر اساس نتایج» بود. لو در سال ۱۸۵۶ با هدایت پارلمان به سوی یک بند اصلی قانون‌گذاری، خود را در تاریخ سرمایه‌داری متمایز ساخت. قانون شرکت‌های سهامی که در سال قبل از آن به تصویب رسیده بود و قانون شرکت‌های با مسئولیت محدود، قانون جدیدی برای شرکت‌ها بر اساس اصل محدودیت محدود ایجاد کردند. لو از اصلاح ساختار تجاری به سوی اصلاح مدارس تحت حمایت دولت حرکت کرد. طرح لو مبتنی بر این فرض بود که «وظیفه دولت در آموزش عمومی، کسب بیشترین میزان ممکن خواندن، نوشتن و حساب برای بیشترین تعداد است.» [۱] بودجه مدارس مطابق عملکرد دانش‌آموزان‌شان در سه مهارت خواندن، نوشتن و حساب تامین می‌شد. هر مدرسه به طور سالانه توسط یک بازرس مورد بازدید قرار می‌گرفت و بازرس از هر دانش‌آموز در زبان انگلیسی و حساب آزمون می‌گرفت. به ازای هر دانش‌آموز و با توجه به موفقیت و شکست او در آزمون، مقداری بودجه دولتی به مدرسه تخصیص می‌یافت. برنامه اصلاحی لو به دنبال کاهش بخشی از هزینه‌ها بود اما فراتر از آن بودجه مدرسه را وابسته به نتایج قابل اندازه‌گیری در مهارت‌های پایه و عملی کرد و با پیوند دادن پرداخت به عملکرد، آموزش را با اصول بازارمحور مطابقت داد. [۲]

برنامه لو توسط متیو آرنولد -منتقد بزرگ فرهنگی- به چالش کشیده شد؛ آرنولد یک بازرس دولت در مدارس بود که لو تصمیم به تغییر آنها داشت. آرنولد دائماً نسبت به گسترش معیارهای مناسب بازار به دیگر حوزه‌های زندگی هشدار می‌داد. او با شجاعت، یک اعتراض عمومی علیه مافوق سیاسی خود به راه انداخت. او در مقاله‌ای با عنوان «برنامه دو بار اصلاح شده»، به مفهوم محدود و مکانیکی آموزش که به طور ضمنی در برنامه به آن اشاره شده بود، حمله کرد. او خاطرنشان کرد که توانایی خواندن هوشمندانه در درجه اول نه با خواندن دروس نسبتاً محدود که

1. Robert Lowe

از یک فرهنگ کلی‌تر، متاثر از خانواده یا محروم بودن از آن یا از محیط مدرسه ناشی می‌شود که یک اشتیاق ذهنی نسبت به خواندن ایجاد می‌کند. بنابراین، هدف مدارس باید «تربیت فکری عمومی» باشد که بدون آن مهارت‌های خواندن و نوشتن رشد نمی‌کنند. [۳] او ابراز تاسف می‌کند که دولت به جای اینکه پاسخگوی «میل شدید طبقات پایین به بالا بردن خودشان» باشد فقط به دنبال تامین مالی ابتدایی‌ترین آموزش‌ها بود. [۴] از آنجا که بسیاری از دانش‌آموزان فقیر هنگام برگزاری آزمون سالانه به طور گریزناپذیری غیبت می‌کنند یا در امتحان ناکام می‌مانند، او پیش‌بینی کرد که اثر اصلی اصلاحات پیشنهادی لو منجر به کاهش بودجه مدارس فقرا می‌شود. او نتیجه گرفت که آموزش افراد قرار بود قربانی «دوستان اقتصاد به هر قیمتی» شود. [۵]

آرنولد دایمما خود را در حال بازدید از مدارسی می‌دید که در آنها دانش‌آموزان، کوه‌هایی از حقایق و حساب‌ها را می‌بلعند، اما از توانایی تحلیل محروم و از درک نشر و نظم پیچیده نیز کاملاً ناتوانند. به آنها نه برای استدلال که برای آمادگی در امتحان آموزش داده می‌شد. [۶] او پیش و به‌ویژه پس از تصویب «پرداخت مبتنی بر عملکرد» انتقاد می‌کرد که چنین آموزشی که عناصر شکل‌دهنده و انسانی بسیاری کمی در آن وجود دارد و مدیران به عنوان نتایج ارزشمند از آن یاد می‌کنند، در حقیقت ماشین‌آلات صرف هستند. [۷] این مفهوم آموزش به‌منزله ماشین‌آلات متناسب با تولید قابل‌اندازه‌گیری خواندن، نوشتن و محاسبه و توانایی پاداش دادن بر مبنای خروجی قابل‌اندازه‌گیری در دهه‌های بعد فروکش کرد تا اینکه در انتهای قرن بیستم منتهی به خیزش موج جدیدی شد.

در هر موج بعدی، ما با منتقدانی همچون آرنولد روبه‌رو خواهیم شد که به هزینه‌های بی‌اندازه پاداش دادن به اندازه‌گیری استاندارد اشاره دارند.

### اندازه‌گیری عملکرد: تیلوریزم

نشانه‌های ثبات معیار در جنبش کارایی مدرسه که در سرتاسر چشم‌انداز آموزشی آمریکا گسترش یافته در دهه ۱۹۱۰ آغاز و تا دهه‌های بعد

ادامه یافت. در سال ۱۹۱۱ سیمون پَتِن<sup>۱</sup>، استاد برجسته مدرسه اقتصاد وارتون درخواست کرد که مدارس مدرکی دال بر مشارکت‌شان در جامعه با نشان دادن نتایجی که «به راحتی دیده و اندازه‌گیری می‌شود» [۸] ارائه دهند. سایر اصلاح‌طلبان به دنبال این بودند که ثمرات جنبش کارایی صنعتی که توسط فردریک وینسلو تیلور<sup>۲</sup> - یک مهندس آمریکایی مبدع اصطلاح «مدیریت علمی» - در سال ۱۹۱۱ ابداع شد را به سیستم مدارس بیاورند. [۹] تیلور تولید آهن خام در کارخانه‌ها را با تجزیه فرآیند به بخش‌های جزئی (از طریق مطالعات زمانی و حرکتی) و تعیین سطوح استاندارد محصول نهایی برای هر شغل تحلیل کرد. کارگرانی که وظایف خود را کندتر از زمان تعیین‌شده انجام دادند، با نرخی پایین‌تر در هر واحد تولیدی به آنها دستمزد پرداخت می‌شد و کسانی که انتظارات را برآورده کرده بودند با نرخ بالاتری پاداش می‌گرفتند. تیلور همچنین از یک سیستم دقیق ارزیابی محل کار و نظارت بر آن حمایت می‌کرد. [۱۰] هدف او افزایش کارایی با استانداردسازی و سرعت بخشیدن به کار در کف کارخانه برای ایجاد تولید انبوه بود.

تخصصی کردن و استانداردسازی وظایف، ضبط و گزارش کلیه فعالیت‌ها، تشویق‌ها و تنبیهات مالی و در یک کلام، همه اینها میراث تیلور و هوادارانش برای نسل‌های بعد بود.

تیلوریسم بر اساس تلاش برای جایگزینی دانش ضمنی کارگران با روش‌های تولید انبوه توسعه‌یافته، برنامه‌ریزی شده، ارزیابی و کنترل‌شده توسط مدیران بود. او نوشت: «تحت مدیریت علمی مدیران فرض می‌کنند ... مسئولیت گردآوری همه دانش سنتی در گذشته و سپس طبقه‌بندی، جدول‌بندی و تقلیل این دانش به قواعد، قوانین و فرمول‌ها بر عهده کارگران بوده است ... بنابراین، همه برنامه‌ریزی‌هایی که در سیستم کهن توسط کارگران انجام می‌شده لازم است که در سیستم نوین توسط مدیریت و در تطابق با قانون علمی انجام شود. [۱۱] بنا بر گفته تیلور «فقط می‌توان از طریق استانداردسازی اجباری

1. Simon Patten

2. Frederick Winslow Taylor

روش‌ها، پذیرش اجباری بهترین تجهیزات و شرایط کار و همکاری اجباری اطمینان یافت که این امر سریع‌تر انجام می‌شود و وظیفه پذیرش اجباری استانداردها و همکاری صرفاً بر عهده مدیریت است.» (ایتالیک‌ها در متن اصلی آمده‌اند) [۱۲]

مضامین تیلوریستی درباره نیاز به کارایی بیشتر از طریق استانداردسازی و نظارت در کتاب درسی بسیار اثرگذار «مدیریت مدرسه دولتی» که در سال ۱۹۱۶ توسط الود. پی. کابری<sup>۱</sup>، رئیس مدرسه آموزش دانشگاه استنفورد منعکس شده بود. [۱۳] مفهوم داوری آموزگاران بر مبنای نمرات امتحانی شاگردان‌شان تا دهه‌ها بعد در جریان بود. یک پژوهشگر آموزش به نام ویلیام لنسلوت<sup>۲</sup> کوشید تا سهم آموزگاران را در یادگیری شاگردان‌شان با آزمودن دانش ریاضی دانش‌آموزان در آغاز و پایان سال تحصیلی برای رسیدن به نمره «تغییر دانش‌آموز» تعیین کند. هرچند مشخص شد برخی آموزگاران موثرتر از برخی دیگر بودند، دستاورد برخی شاگردان که توسط بهترین آموزگاران تعلیم می‌دیدند، کاملاً متوسط بود. [۱۴] در اوایل قرن بیست‌ویکم، همین مفهوم با عنوان «امتیاز ارزش افزوده» و سپس در سال‌های ریاست جمهوری اوباما با نام «رشد دانش‌آموز» احیا شد. [۱۵]

شیوه‌های تیلوریستی سامان‌دهی تولید کارخانه، به طور فزاینده‌ای در طیف گسترده‌ای از صنایع تولیدی در دوره میان جنگ مورد پذیرش قرار گرفت. تا دهه ۱۹۵۰ این شیوه‌ها در شرکت‌هایی همچون جنرال موتور یک قاعده بودند، همان‌گونه که جامعه‌شناسی به نام دنیل بل<sup>۳</sup> خاطرنشان می‌کند: «روبنای مدیریتی که تولید را سامان‌دهی و هدایت می‌کند ... تمام کارهای فکری ممکن را از فروشگاه دور می‌کند، همه چیز در مرکز و در دپارتمان‌های برنامه‌ریزی و زمان‌بندی و طراحی انجام می‌شود.» نتیجه این امر بی‌حسی در کارگران پایین سلسله‌مراتب را تقویت کرد. [۱۶] در پایان قرن معیارها این شیوه‌های سازمانی را از بخش تولید به بخش خدمات آوردند.

1. Ellwood.P.Cubberley

2. William Lancelot

3. Daniel Bell

## مدیریت‌گرایی و اندازه‌گیری

تیلوریسم توسط مهندسان توسعه یافت، اما نقش دیگر در فرهنگ پاسخگویی را اندازه‌گیری استانداردسازی شده ایفا می‌کند که از حرفه حسابداری آمده است. رابرت مک‌نامارا<sup>۱</sup> حسابداری بود که در سن ۲۴ سالگی به جوان‌ترین استاد مدرسه اقتصاد هاروارد بدل شد و پیام معیارها را به بزرگ‌ترین سازمان ایالات متحده یعنی ارتش آمریکا برد. دهه‌هایی که مک‌نامارا در حکم استاد مدرسه اقتصاد ظهور کرد، از مدیرعامل شرکت فورد موتور<sup>۲</sup> تا وزارت دفاع و در نهایت رئیس بانک جهانی، همگی شاهد تغییر و تحول در مدارس اقتصاد آمریکا بودند. در دوره اولیه مدارس اقتصاد بر آمادگی دانشجویان‌شان برای مشاغل در صنایع یا شرکت‌ها تمرکز داشتند. از دهه ۵۰ به بعد، مدرسه اقتصاد ایده‌آل به مدیر عمومی که با مجموعه‌ای از مهارت‌ها تجهیز شده و در صنایع خاص مستقل است، تبدیل شد.

اکنون هسته تخصص مدیریتی به عنوان مجموعه‌ای متمایز از مهارت‌ها و تکنیک‌ها تعریف می‌شد که بر تسلط به روش‌شناسی‌های کمی تمرکز داشت. [۱۷] تصمیمات مبتنی بر اعداد، علمی به نظر می‌رسید زیرا تصور می‌شد که اعداد بر عینیت و دقت دلالت دارند. [۱۸] نظریه‌پردازان و معلمان مدیریت کسانی بودند که این خرد جدید را در اداره‌ها پیش بردند، کسانی که شلی<sup>۳</sup> یک بار آنان را همچون شاعران، در حکم «قانون‌گذاران ناشناخته بشر» توصیف کرد. [۱۹]

پیش از آن، «تخصص» به معنای انباشت دانش شغلی بلندمدت در یک حوزه خاص بود، مانند پیشرفت پله به پله در یک نهاد یا کسب‌وکار مشابه؛ انباشت آنچه که اقتصاددانان آن را «فوت‌وفن وظیفه تخصصی»<sup>۴</sup> می‌نامند. مدیران خودرو که «دست‌اندرکاران اتومبیل» بودند، کسانی که بیشتر زندگی حرفه‌ای خود را در صنعت خودروسازی سپری کرده بودند.

1. Robert McNamara

2. Ford Motor Company

3. Shelley

4. Task-specific know-how

آنها به طور روزافزونی با افرادی همچون مک‌نامارا و «شمارندگان لوبیا»<sup>۱</sup> که در محاسبات سود و زیان ماهر بودند، جایگزین شدند. [۲۰]

با گذشت زمان، این تلاش برای تبدیل مدیریت به یک علم برای آماده‌سازی علاقه‌مندان به مناصب اجرایی در شرکت‌های آمریکا وارد آموزه‌های مدیریت‌گرایی شد. بر این اساس، نقش قضاوت مبتنی بر تجربه و دانش عمیق در یک زمینه، تقلیل یافت. فرض مدیریت‌گرایی این است که تفاوت‌های میان سازمان‌ها - شرکت‌های خصوصی، سازمان‌های دولتی و دانشگاه‌ها- در مقایسه با شباهت‌ها از اهمیت کمتری برخوردارند. از این رو عملکرد همه سازمان‌ها را می‌توان با ابزار مشابه تکنیک‌ها و مهارت‌های مدیریتی بهینه‌سازی کرد. [۲۱] چه بسا به قضاوت و تخصص مبتنی بر تجربه به عنوان تسهیل‌گری بیندیشیم که با ارائه فوت و فن وظیفه تخصصی موجب شکوفایی سازمان‌ها می‌شوند. مدیریت‌گرایی تحت لوای معیارها اگر به همه چیز به دیده تحقیر نگاه کند اما گرایش به نادیده گرفتن دارد.

مانند وزیر دفاع که مسئول پیگیری جنگ در ویتنام بود، مک‌نامارا نیز از معیار «شمارش اجساد»<sup>۲</sup> به منزله یک شاخص قابل اطمینان پیشرفت آمریکایی‌ها به عنوان پیروزی در جنگ حمایت کرد. با این وجود فقط تعداد اندکی از فرماندهان این حوزه، تعداد اجساد را در حکم یک معیار صحیح و معتبر موفقیت پذیرفتند و بسیاری از آنان می‌دانستند که این تعداد یا اغراق است یا کاملاً ساختگی. [۲۲] در نتیجه، در فرمول‌بندی مختصر و مفید کینت کوکی‌یر<sup>۳</sup> و ویکتور مایر شونبرگر<sup>۴</sup> «مغالطه کمی»<sup>۵</sup> وجود داشت. [۲۳]

پنتاگون مک‌نامارا با آنچه که استراتژیست نظامی ادوارد لوتواک<sup>۶</sup>

1. Bean Counter

فردی که غالباً حسابدار است و به طور فزاینده‌ای دائماً در حال حساب و کتاب و کنترل هزینه‌ها است. این عبارت به معنای جزئی‌نگر نیز هست. م

2. Body Counts

3. Kenneth Cukier

4. Viktor Mayer-Schönberger

5. Quagmire of Quantification

6. Edward Luttwak

آن را «جایگزین عمده تجزیه و تحلیل افراد غیرنظامی برای تخصص نظامی» می‌نامید، مشخص می‌شد. نسل جدید «تحلیلگران سیستم»، استانداردهای جدید نظم فکری را معرفی کرد و به طور گسترده‌ای موجب بهبود روش‌های سامان‌دهی و حسابداری شد اما در عین حال یک ناتوانی در فهم مهم‌ترین جوانب قدرت نظامی را که غیرقابل اندازه‌گیری است، نیز آموزش داد. [۲۴] نیروهای مختلف ارتش به دنبال به حداکثر رساندن «تولید» قابل اندازه‌گیری بودند: نیروی هوایی از طریق تعداد انواع بمب‌گذاری‌ها، توپخانه از طریق تعداد گلوله‌های شلیک شده و پیاده نظام از طریق شمارش اجساد منعکس‌کننده شاخص‌های آماری ابداع‌شده توسط مکنامارا و همکارانش در پنتاگون بودند. اما همان‌گونه که لوتواک می‌نویسد: «در جنگ بدون جبهه، جایی که روی نقشه هیچ خطوط مشخصی برای نشان دادن پیروزی و شکست وجود ندارد، یگانه معیار حقیقی پیشرفت باید سیاسی و غیرقابل اندازه‌گیری باشد: تدبیر بر خواست دشمن برای ادامه جنگ.» [۲۵]

انتقاد لوتواک از تاسیس ارتش آمریکا که در سال ۱۹۸۴ به چاپ رسید بر این واقعیت متمرکز بود که اخلاق مدیریتی، هم در رهبری غیرنظامیان و هم در رهبری نظامیان نفوذ کرده و موجب دنبال کردن «کارایی‌های» قابل اندازه‌گیری شده است که با نوع تفکر استراتژیک مورد نیاز ارتش مغایرت دارد. «تحت هدایت مقامات غیرنظامی - که بسیاری از آنها توجه اندکی به ناآگاهی خود از استراتژی، مهارت عملیاتی و تاکتیک‌ها دارند و خود را به عنوان مدیرانی معرفی می‌کنند که توانایی مدیریت همه امور را فارغ از محتوای آنها دارند - تشکیلات ارتش، خود از مدت‌ها قبل به دنبال رسیدن به کارایی تجاری به منزله هدفی غایی بود.» مقامات نظامی نیز به طور فزاینده‌ای این چشم‌انداز مدیریتی را پذیرفته و به دنبال گرفتن مدرک در رشته‌های مدیریت تجاری، مدیریت یا اقتصاد بودند. این امر به چیزی انجامید که لوتواک آنرا «تعصب مادی‌گرا»<sup>۱</sup> می‌نامید و موجب اندازه‌گیری ورودی‌ها

و خروجی‌های مشهود (مانند قدرت شلیک) شد و نه عوامل انسانی نامشهود مانند استراتژی، رهبری، انسجام گروهی و روحیه سربازان. [۲۶] آنچه به دقت اندازه‌گیری می‌شد تمایل داشت تا بر آنچه به راستی مهم بود، سایه بیندازد. او یادآور شد: «در حالی که ورودی‌های مادی همگی حقایقی قابل استناد هستند و هزینه‌ها نیز به طور دقیق به دلار و سنت مشخص شده‌اند، موارد نامشهود حتی به سختی تعریف و تعیین می‌شوند و کلاً غیرقابل اندازه‌گیری هستند.» [۲۷]

خواه توصیفات لوتواک کاملاً منصفانه باشد یا نه، بیشتر انتقاد او به تاسیسات نظامی آمریکا درباره انتقال طیف گسترده‌ای از نهادها در درون ایالات متحده آمریکا و فراتر از آن بود.

یکی از پیش‌برندگان ثبات معیار، ظهور مشاوران مدیریتی بود که مجهز به مهارت‌های تجزیه و تحلیل کمی و مطابق با این اصل بودند: «اگر نمی‌توانی اندازه‌گیری کنی، نمی‌توانی مدیریت کنی.» [۲۸] اتکا به اعداد و داده‌های دستکاری‌شده کمی نه تنها بر تخصص علمی مبتنی بر شواهد «محکم» اثرگذار است، بلکه نیاز نهادها به دانش ویژه و جزئی افرادی را که مشاوره به آنها می‌فروشد نیز به حداقل می‌رساند. [۲۹] فرهنگ مدیریت خواستار داده‌های بیشتر - داده‌های استاندارد و عددی - است.

## فصل ۴

## چرا معیارها بسیار محبوب شدند

همان‌گونه که در مطالعات موردی خود خواهیم دید و در فصل پایانی مفصل‌تر جست‌وجو خواهیم کرد، موقعیت‌هایی وجود دارد که در آن معیارها به خوبی کار می‌کنند اما موقعیت‌های بسیاری نیز وجود دارد که معیار پاسخگویی، کژکارکردی بیشتری دارد تا اینکه کارکردی باشد. پس چگونه باید فاصله میان اثربخشی فرهنگ اندازه‌گیری، پاسخگویی، شفافیت و فراگیر بودن آن فرهنگ را حساب کنیم؟ و با وجود همه اشکالات چرا باز هم معیارها از محبوبیت زیادی برخوردارند؟ در حالی که پاسخ واحدی وجود ندارد اما اثبات آن نیز چندان دشوار نیست. در اینجا به ارائه چند حدس آگاهانه می‌پردازیم.

## بی‌اعتمادی به قضاوت

با کاهش اعتماد، تقاضا برای پاسخگویی و شفافیت اندازه‌گیری شده افزایش می‌یابد. میان یک جامعه دموکراتیک با بسیج اجتماعی قابل توجه و ناهمگونی قومی بیشتر و فرهنگ پاسخگویی اندازه‌گیری شده، یک

پیوستگی انتخابی وجود دارد. در جوامعی با یک طبقه بالای بین‌نسلی جاافتاده، به احتمال زیاد اعضای این طبقه بیشتر در پی احساس تضمین و امنیت نسبت به موقعیت و جایگاه خود هستند، به یکدیگر اعتماد دارند، میزانی از دانش ضمنی را برای اداره کردن خانواده‌های‌شان دارند و تا اندازه‌ای دارای اعتماد به نفس در قضاوت‌های‌شان هستند (چه این اعتماد به نفس قابل توجیه باشد یا نباشد). [۱] برعکس، در جوامع شایسته‌سالار با نخبگانی صریح‌تر و قابل‌تغییرتر، افرادی که به قدرت می‌رسند در قضاوت‌های‌شان کمتر احساس تضمین و امنیت می‌کنند و احتمال اینکه به دنبال معیارهای عینی برای تصمیم‌گیری باشند، بیشتر است. [۲] اعداد، جریان عینیت را منتقل می‌کنند و بر نبود قضاوت ذهنی دلالت دارند. اعداد، «سخت و محکم» شناخته می‌شوند و به همین خاطر خطر کمتری برای افرادی دارند که نسبت به قضاوت خود مردد هستند.

همچنین معیارهای عددی، ظاهر پاسخگویی و شفافیت به خود می‌گیرند (البته اگر کسی منشأ و ارتباط آنها را بررسی نکند). بخش خوب جذابیت آنها این است که برای همه به راحتی قابل درکند؛ همان‌گونه که استفان کالینی<sup>۱</sup> دانشجوی ادبیات کمبریج مشاهده کرده است: «بحث عمومی در دموکراسی‌های لیبرال مدرن ترکیب ارزش‌های فایده‌گرا با بی‌اعتمادی به رویه‌هایی است که به طور مکانیکی قابلیت جهان‌شمولی ندارند.» [۳]

جست‌وجو برای معیارهای عددی پاسخگویی در فرهنگ‌هایی با اعتماد اجتماعی پایین از جذابیت ویژه‌ای برخوردار است و بی‌اعتمادی به قدرت به منزله یک موضوع مهم تکراری در فرهنگ آمریکایی از دهه ۱۹۶۰ وجود داشته است. بنابراین در سیاست، اجرا و بسیاری از حوزه‌های دیگر، اعداد به طور دقیق ارزش‌گذاری شده‌اند زیرا اعتماد را به قضاوت‌های ذهنی و مبتنی بر تجربه افراد در مسند قدرت بازمی‌گردانند. جست‌وجوی معیارهای پاسخگویی برای راست و چپ سیاسی جذاب است. پیوستگی

1. Stefan Collini

زیادی میان این معیار و پوپولیست، برابری خواه مشکوک به قدرت مبتنی بر طبقه و تخصص و پیشینه وجود دارد.

تقاضا برای «پاسخگویی» بیشتر که بازتاب آن را در گوگل انگرام دیدیم، از بی‌اعتمادی روزافزون نهادها و نارضایتی از قدرت مبتنی بر تخصص در ایالات متحده آمریکا (و به طور قابل توجهی دیگر جوامع غربی) از دهه ۱۹۶۰ به بعد تغذیه می‌شود. جرج برنارد شاو<sup>۱</sup> در نمایشنامه خود با عنوان «معضل پزشک<sup>۲</sup>» نوشته است: «هر حرفه‌ای، توطئه علیه عوام است.» با آغاز دهه ۱۹۷۰، آنچه برای شاو یک شوخی به حساب می‌آمد، بدل به یک فرض کاربردی در سیاست عمومی شد. چپ و راست به دنبال معیارها بودند؛ هرچند نه با دلایل یکسان. ظن به قدرت ذاتا در چپ سیاسی پس از ۱۹۶۰ وجود داشت: اتکای به قضاوت متخصصان، تسلیم پیش‌داری‌های نخبگان مستقر بود. بنابراین، چپ دلایل خود را برای پیشبرد یک برنامه کاری داشت که مدعایش ایجاد پاسخگویی و شفافیت در نهادها و استفاده ظاهری از استانداردهای عینی و علمی در عملکرد اندازه‌گیری شده بود.

در راست سیاسی نیز این ظن وجود داشت و گاهی مستدل بود که نهادهای بخش عمومی بیشتر در راستای منافع کارمندان خود اداره می‌شوند تا مشتریان و موسسان. در برخی مدارس، بخش‌های پلیس و دیگر ادارات دولتی، وقت‌گشی<sup>۳</sup> یک واقعیت است، حتی اگر آنچنان که منتقدان آن را مرسوم و جهانی می‌دانند، نباشد. فرهنگ پاسخگویی مبتنی بر معیار سنجش یک تلاش قابل فهم برای شکستن سنگر حکومت سالخوردگان بود. هنگامی که موسسات نهادی تحت حمله پوپولیست قرار می‌گیرند، برای نشان دادن اثربخشی خود به معیارها به عنوان ابزارهای دفاعی پناه می‌برند.

در یک دور باطل، نبود اعتماد اجتماعی منجر به ستایش اغراق‌آمیز معیارها و ایمان به معیارها موجب کاهش اعتماد به قضاوت می‌شود. فیلیپ کی. هاوارد<sup>۴</sup> در یک مجموعه کتاب این بحث را مطرح کرده است

1. George Bernard Shaw

2. Doctor's Dilemma

3. Time-serving

4. Philip K. Howard

که کاهش اعتماد به طرز فکر جدیدی می‌انجامد که در این طرز فکر «بی‌اعتباری انتخاب انسان تنها یک نظریه نیست ... بلکه یک حکمت الهی است... انتخاب انسان بسیار خطرناک تلقی می‌شود.» و در نتیجه «مقامات دیگر مجاز نیستند تا بر اساس بهترین قضاوت خود عمل کنند» [۴] یا «نمی‌توانند در موقعیت ویژه‌ای که نیازمند قضاوت است، اعمال نظر کنند.» [۵] نتیجه این امر، اعمال قوانین و مقررات بیش از حد است: شبکه محکم‌تری از قوانین شامل تکثیر قوانین در سازمان‌ها. [۶] اغلب، معیارها ابزارهایی برای محکم‌تر کردن شبکه نیز فراهم می‌کنند. اندازه‌گیری بیش از حد نوعی اعمال قانون بیش از حد است، همان‌گونه که اندازه‌گیری نادرست شکلی از اعمال قانون نادرست است.

دیگر انگیزه اندازه‌گیری عملکرد، ترس از دعوی قضایی به عنوان نتیجه گسترش مسئولیت در قانون مجازات آمریکا است. در دوره قرن بیستم، موانع آموزه اولیه علیه تحت تعقیب قرار دادن پزشکان، بیمارستان‌ها، تولیدکنندگان و شهرداری‌ها از بین رفت. گسترش حقوق مدنی و قانون محیط‌زیست، دعاوی قضایی بیشتر را تشویق و ترویج کرد. [۷] در حوزه اشتغال، قوانین حقوق مدنی، موانع بیشتری از ثبت و ضبط و مقررات دست‌وپاگیر در شرکت‌های خصوصی و سازمان‌های دولتی ایجاد کرد. [۸] در نتیجه: هزینه برای وکلا بیشتر و بیشتر شد. و شناخت ایالات متحده به منزله یک جامعه قابل‌پیگرد [۹] موجب ایجاد نگرانی در تحت تعقیب قرار گرفتن شد که این امر خود به دفاعی بودن و خطرگریزی می‌انجامد. اصرار بر مستندسازی هر تصمیم به عینی‌ترین شیوه ممکن، به‌گونه‌ای که موجب شود تصمیمات مربوط به استخدام و ارتقا برای مراجع نظارتی شفاف باشد یا در دعاوی قضایی مورد استفاده قرار گیرد، انگیزه دیگری برای اندازه‌گیری عملکرد است.

### انتقاد از مشاغل و ستایش اغراق آمیز انتخاب

در راست سیاسی، بی‌اعتمادی به نهادهای بخش عمومی منجر به اظهار این عقیده شد که مشکل با بخش‌های غیرانتفاعی (حکومت، مدارس،

دانشگاه‌ها) این است که آنها «خط پایانی ندارند» و بنابراین روشی برای محاسبه موفقیت و شکست‌شان نیز وجود ندارد. در راستای این روش تفکر، راه حل عبارت است از ایجاد یک خط پایانی جایگزین به شکل «عینی» - و ترجیحا عددی - که رویه‌های استاندارد را اندازه بگیرد. یک روند موازی از جانب طرفداران سلامت زنان و جنبش‌های پس از آن است که نهادهای تاسیس‌شده (مانند پزشکان) را به چالش می‌کشد و به دنبال این است که آنها را پاسخگوتر کند. این طرفداران به دنبال نظارت بیشتر بر مراقبت‌های پزشکی ارائه‌شده به بیماران بودند. این امر مستلزم دادن انتخاب و اطلاعات بیشتر - شامل معیارهای عملکرد - به بیماران در راستای آگاه ساختن آنان از انتخاب‌های‌شان است. جاده توانمندسازی با معیارها هموار می‌شود.

در هر حوزه یکی پس از دیگری، معرفی اندازه‌گیری بیشتر به نام پاسخگویی، چراغی را روی مشکلات واقعی تاباند که شامل متغیرهایی در فعالیت حرفه‌ای می‌شد که احتمالاً ریشه در «علم» دارد و همچنین شکاف‌هایی را در عملکرد که پیشتر دیده نمی‌شدند یا غیرمستند بودند، در بر می‌گرفت. تاثیر این افشاگری‌ها هم موجب کم شدن اعتقاد به قضاوت حرفه‌ای و هم موجب ایجاد فشار برای یافتن راه‌های جدید شد؛ راه‌هایی که مستلزم اندازه‌گیری بیشتر در راستای نظارت بر متخصصانی شد که اخلاق حرفه‌ای آنها مورد تردید قرار گرفته بود.

این روندها ارتباط نزدیکی با افزایش نفوذ ایدئولوژی انتخاب مصرف‌کننده - باور به اینکه پس از ارائه اطلاعات، افراد در مراقبت پزشکی، آموزش، بازنشستگی، برنامه‌ریزی و ... انتخاب درستی انجام می‌دهند - دارند. در حقیقت، اغلب اوقات افراد توانایی تصمیم‌گیری درباره بهترین ارائه‌دهنده خدمات را دارند اما نه همیشه؛ در برخی حوزه‌ها، انتخاب کاملاً هیجانی است. برای مثال در مراقبت‌های بهداشتی، انتخاب‌های مربوط به پزشکان یا بیمارستان‌ها زمانی صورت می‌گیرند که یا افراد سالمند و تمایلی به نگرانی و درگیر شدن در امور پزشکی را ندارند یا بیمارند و از این رو نسبت به تصمیمات خود مضطرب می‌شوند که در این

صورت توانایی‌شان در پردازش معیارهای پیچیده و اغلب متضاد کاهش می‌یابد. با این حال تا دهه ۱۹۹۰ با وجود مطالعاتی که نشان‌دهنده توانمندسازی بیمار بودند، کاری در راستای مهار هزینه‌ها یا بهبود کیفیت مراقبت انجام نشد و حتی الگوی بیمار در حکم مصرف‌کننده در بازار خدمات پزشکی در میان سیاست‌مداران و سیاست‌گذاران هر دو طیف سیاسی محبوبیت بیشتری پیدا کرد.<sup>[۱۰]</sup>

### بیماری هزینه<sup>۱</sup>

دیگر انگیزه پاسخگویی در حوزه‌های پزشکی و آموزش ناشی از این واقعیت است که هزینه نسبی این خدمات در مقایسه با هزینه‌های بیشتر کالاهای مصرفی افزایش یافته است. بخشی از این دلیل در «بیماری هزینه» نهفته است؛ پدیده‌ای که نخستین بار توسط دو اقتصاددان به نام‌های ویلیام بامول<sup>۲</sup> و ویلیام بوئن<sup>۳</sup> در سال ۱۹۶۶ شناسایی شد. آن دو شاهد افزایش مداوم بهره‌وری در تولید طی صد سال گذشته بودند و آن را به طور عمده نتیجه فناوری پیشرفته می‌دانستند.<sup>[۱۱]</sup> از آنجا که پیشرفت‌های فناوری و افزایش تجارت جهانی منجر به هزینه‌های رو به کاهش بیشتر کالاهای مصرفی شده است، هزینه‌های نسبتاً بالای پزشکی، آموزش و خدمات انسانی مشابه بیش از پیش به چشم می‌آید و بر نارضایتی عمومی تمرکز دارد. با گذشت سال‌ها، این روندها منجر به افزایش فشار عمومی برای کارایی و پاسخگویی بیشتر شده است - با وجود دشواری در اندازه‌گیری (سنجش) ورودی‌ها، خروجی‌ها و دست‌آخر، بهره‌وری در این دو حوزه.<sup>[۱۲]</sup> باید افزود که پیشرفت‌ها در فناوری پزشکی و داروهای موثرتر شاید باز هم به طور مشروع و قانونی بر این هزینه‌ها بیفزاید: اگر افراد مدت زمان بیشتری عمر کنند یا زندگی دلپذیرتری داشته باشند و زمان کمتری را در بیمارستان سپری کنند، آن وقت این هزینه‌ها می‌تواند ارزشش را داشته باشد.

1. The Cost Disease

2. William Baumol

3. William Bowen

## رهبری در بحبوحه پیچیدگی سازمانی

سایر نیروهای اقتصادی نیز در ایجاد فشار برای اندازه‌گیری‌های کمی و قابل سنجش نقش دارند. در حالی که سازمان‌ها (شرکت‌ها، دانشگاه‌ها و موسسات دولتی) بزرگ‌تر و متنوع‌تر می‌شوند، فاصله بیشتری میان مدیریت ارشد و افراد در رده‌های سازمانی پایین‌تر ایجاد می‌شود که در زنجیره سازمانی مشغول انجام فعالیت‌های واقعی هستند که سازمان متعهد به انجام آنهاست. هنگامی که نهادها بسیار بزرگ، پیچیده و دارای بخش‌های متفاوتند، درک آنها به‌سادگی ممکن نیست. کسانی که در موقعیت بالاتری قرار دارند، محدودیت شناختی بیشتری نسبت به ما دارند: تصمیم‌گیری با وجود محدودیت زمانی و توانایی روبرویی با اطلاعات بیش از حد. معیارها، ابزارهایی و سوسه‌انگیز برای مقابله با «عقلانیت محدود» و درگیر شدن با موضوعاتی هستند که فراتر از درک فرد است. تصور کنید برای نمونه، رئیس یک دانشگاه بزرگ، شرکت یا پارلمان هستید. البته ممکن است به نظرات آگاهانه زبردستان باتجربه خود اعتماد کنید اما به احتمال زیاد آنها تمایل ذاتی به حفظ وضع موجود دارند: گفته شاعر و تاریخ‌نگار مرحوم، رابرت کانکوئست<sup>۱</sup> را به خاطر بسپارید: «هرکس درباره آنچه که بهتر می‌داند، محافظه‌کار است.» اما اگر بخواهید پویایی یا تغییری در سازمانی که ریاست شما مد نظر دارد ایجاد کنید (که این یک و سوسه معمول برای وزرای کابینه، روسای دانشگاه و مدیران عامل شرکت‌هایی است که به دنبال «کاری مهم و اثرگذار» هستند) چه؟ در این صورت به نظر می‌رسد یافتن و دسترسی به «اعداد» مستقیم‌ترین میان‌بر برای درک سازمانی شما باشد.

مشکل این است که تلاش مدیریت برای درک یک مجموعه سازمانی اغلب منجر می‌شود به چیزی که ایو موریو<sup>۲</sup> و پیترو تولمن<sup>۳</sup> به آن «پیچیدگی»<sup>۴</sup> می‌گویند: گسترش رویه‌های گزارش‌دهی و تصمیم‌گیری

1. Robert Conquest

2. Yves Morieux

3. Peter Tollman

4. Complicatedness

که نیازمند مجموعه‌های هماهنگ، جلسه‌ها و گزارش‌نویسی بیشتر است. با زمان‌هایی که صرف گزارش، جلسه و هماهنگی می‌شود، زمانی برای انجام فعالیت باقی نمی‌ماند. [۱۳]

این از بین رفتن زمان و تلاش با گرایش روسا به ثبات معیارها به جای قضاوت تجربی افرادی که برای آنها کار می‌کنند، تشدید می‌شود. آنها خواهان تلاش برای کنترل بیشتر زیردستان خود از طریق استراتژی‌های مختلف هستند که معیارها نیز جزء اصلی آنها به شمار می‌رود. تقاضا برای جریان دائم گزارش و داده‌های استاندارد -خواسته یا ناخواسته- بر میزان استقلال افراد در رده‌های پایین سلسله‌مراتب سازمانی اثرگذار است و موجب کاهش آن می‌شود -و افرادی که نسبت به نوآوری‌های مبتنی بر معیارها تردید دارند با عناوین غیرمنطقی، یا «نفع‌شخصی‌طلب»<sup>۱</sup> یا مقاومت در برابر تغییر نادیده گرفته می‌شوند. بعضی بوروکراسی‌های آمریکایی (شرکتی و دولتی و غیرانتفاعی) یکسری ویژگی‌های فرهنگی دارند که بر اساس آن فرض بر این است که هر فرد می‌تواند و باید از طریق سلسله‌مراتب صعودی پست‌ها، هم در درون سازمان‌ها و هم میان سازمان‌ها، در چرخش باشد. این امر برخلاف توسعه میزان تخصص است که امکان ارزیابی معناداری از مفاد و اهمیت کیفی کار انجام‌شده توسط زیردستان را مهیا می‌کند و به اینسان اتکا به معیارهای قابل اندازه‌گیری و کمی جذاب است.

مدیران عامل، روسای دانشگاه و روسای سازمان‌های دولتی اکنون نسبت به گذشته بیشتر از یک سازمان به سازمان دیگر می‌روند. یک کیمیاگر بیگانه برابری‌خواه غالباً فرض می‌کند که باید فرد بهتری وجود داشته باشد که در خارج از سازمان پیدا می‌شود تا در داخل سازمان: هیچ‌کس درون سازمان به اندازه کافی برای ترقی کردن خوب نیست اما افراد ناشناخته از جاهای دیگر شاید بهتر باشند. [۱۴] این فرض به تغییر و گردش رهبران، روسا و مدیران رده بالایی می‌انجامد که با دانش محدودی به پست‌های جدید خود در نهادهایی رسیده‌اند که

1. Self-interested

مدیریت آنها را بر عهده دارند. بنابراین آنها اتکای بیشتری به معیارها -ترجیحا معیارهایی که در سازمانها بیشتر به یکدیگر شبیه است (با نام «بهترین راهکار»)- دارند. این افراد غیر خودی که به خودی بدل شده‌اند، دانش زمینه‌ای عمیق مبتنی بر تجربه را ندارند و بیشتر به شیوه‌های استاندارد اندازه‌گیری وابسته هستند. نه تنها این امر، بلکه با توجه به خروج نهایی آنها برای یک شغل جدید در سازمانی دیگر، مدیران ستار همواره مراقب معیارهای عملکرد هستند تا بتوانند هنگام تماس کارآفرینان، در پست‌ها مستقر شوند.

### جذابیت معیارها

عامل دیگر، گسترش فناوری اطلاعات (IT) است. در ابتدای دهه ۱۹۸۰، اختراع و پذیرش سریع صفحه‌های گسترده<sup>۱</sup> الکترونیکی موجب تسهیل جدول‌بندی، ترسیم آماری و دستکاری اشکال و جداول شد که اثرات بسیار گسترده‌ای داشت. همان‌گونه که استیون لوی<sup>۲</sup> به عنوان یک تحلیلگر دوراندیش این پدیده، در سال ۱۹۸۴ نوشته است:

«صفحه گسترده یک ابزار است -اما در عین حال یک نگرش نیز هست- بیان واقعیت توسط اعداد؛ ... از آنجا که صفحه گسترده می‌تواند کارهای بسیار مهمی انجام دهد، استفاده‌کنندگان از آن تمایل دارند تا از این واقعیت مهم چشم‌پوشی کنند که کسب‌وکارهای خیالی که آنها می‌توانند روی کامپیوترهای خود ایجاد کنند، صرفا یک خیالند. شما واقعا نمی‌توانید یک کسب‌وکار را درون یک کامپیوتر کپی کنید؛ تنها ممکن است برخی جنبه‌های آن را بتوانید. و از آنجا که اعداد نقاط قوت صفحات گسترده است، جنبه‌هایی مورد تاکید قرار می‌گیرد که به‌سادگی در اعداد تجسم یابد و عوامل نامشهود، به‌سادگی قابل سنجش نخواهد بود.» [۱۵]

1. Spreadsheet

صفحه گسترده - جدولی از اعداد است که به صورت سطرها و ستون‌ها مرتب شده است. صفحه‌های گسترده کاغذی، قبل از اختراع کامپیوترها برای عملیات تجاری به کار می‌رفتند. نخستین صفحه گسترده کامپیوتری VisiCalc بود که در پایان دهه ۷۰ روی کامپیوتر Apple II اجرا می‌شد. مشهورترین برنامه‌های صفحه گسترده امروزی Lotus 1-2-3 و Excel است. م

2. Steven Levy

سِتِ کِلارَمَن که از موفق‌ترین سرمایه‌گذاران ارزش در نسل خود است، با این نظر موافق بود، او در سال ۱۹۹۱ اخطار داد که صفحات گسترده موجب توهم عمق تحلیل می‌شود. [۱۶]

از آن زمان به بعد، فرصت‌های رو به رشد جمع‌آوری داده‌ها و کاهش هزینه‌های انجام این کار به این الگوی رفتاری کمک می‌کند که مطابق آن، داده‌ها پاسخند؛ در حالی که این پرسش‌ها هستند که سازمان‌ها باید به آن بپردازند. اغلب این باور نسنجیده وجود دارد که گردآوری داده‌ها و به اشتراک گذاشتن آنها به طور گسترده در درون سازمان منجر به پیشرفت‌هایی خواهد شد حتی اگر اطلاعات زیاد برای تبدیل شدن به داده‌ای که به‌سادگی منتقل می‌شود، باید از تفاوت‌های ظریف و زمینه‌ای عبور کند.

## فصل ۵

## کارفرمایان، کارگزاران و انگیزه‌ها

در همان دهه‌هایی که تخصص حرفه‌ای مورد انتقاد قرار می‌گرفت، شرکت‌های تجاری نیز به دلیل طرفداری از منافع مدیران‌شان در برابر سهامداران خود مورد انتقاد قرار گرفتند.

این مفهوم در دهه ۱۹۷۰ محبوب شد و به اصل دانشگاهی «نظریه کارفرما-کارگزار» انجامید. [۱] نسخه اولیه نظریه در ادبیات مدیریت توجهات را به سوی شکاف میان اهداف تعیین‌شده نهادها و افرادی که آنها را اجرا می‌کنند و افرادی که به وسیله این مجریان استخدام شده‌اند، جلب کرد. این نظریه بر همسویی منافع سهامداران در حداکثر سودآوری و قیمت سهام با منافع مدیران شرکت‌ها تمرکز دارد و اینکه اولویت‌های آنها چه‌بسا با اهداف‌شان همسو نباشد.

نظریه کارفرما-کارگزار به طور انتزاعی بر این باور است که افراد استخدام‌شده در نهادها قابل اعتماد نیستند و فعالیت‌های آنها باید ارزیابی و اندازه‌گیری شود؛ افرادی که کار اندازه‌گیری را انجام می‌دهند باید نسبت به افرادی که دانش تجربی درباره نهادها ندارند شفاف باشند و

پاداش‌ها و تنبیهات مالی نیز موثرترین روش برای تشویق «کارگزاران» است. [۲] همچنین در اینجا اعداد به عنوان ضمانت عینیت و جایگزینی برای دانش فراگیر و اعتماد شخصی دیده می‌شود. [۳]

نظریه کارفرما-کارگزار در ابتدا طرح‌هایی برای پاداش مدیران عامل بر اساس قیمت سهام و سود شرکت‌های تحت اداره‌شان ارائه داد. بعدها این نظریه در تدارک این بود که برای مدیران ارشد اختیاراتی در خصوص سهام شرکت‌های‌شان فراهم کند. این ایده در هر مورد این بود که انگیزه‌های مدیران را با صاحبان بنگاه همسو سازد که فرض می‌شد یگانه منفعت‌شان به شکلی کاملاً منطقی، سودآوری شرکت است.

نظریه کارفرما-کارگزار، سازمان‌ها را به منزله شبکه‌هایی از روابط، میان افرادی تصور می‌کند که منفعتی (کارفرما) دارند و افرادی که استخدام شده‌اند تا این منفعت را دنبال و به مرحله عمل برسانند (کارگزار). چشم‌انداز و پیش‌فرض این است که منافع کارفرمایان با کارگزاران از یکدیگر فاصله می‌گیرد. مثلاً ممکن است منافع سهامداران یک شرکت به حداکثر رساندن سود و بازگشت سرمایه باشد اما شاید منافع مدیران شرکت این باشد که دفاتر پرزرق‌وبرق و هواپیماهای خصوصی مجلل برای خود داشته باشند تا به این وسیله موقعیت‌های خود را گسترش دهند و همچنین ممکن است منافع کارکنان سطوح پایین‌تر، مطالبه حقوق و حداقل‌سازی میزان کار باشد. چالش اصلی کارفرمایان این است که برای کارگزاران انگیزه‌هایی ایجاد کنند که آنها اولویت‌های کارفرما را به اولویت‌های خود ترجیح دهند. یک مشکل جانبی برای کارفرما مساله نظارت است: چگونه می‌توان فهمید که کارگزاران به‌راستی مشغول کارند و اهداف را چقدر خوب محقق می‌سازند؟ از این رو، وظایف دوگانه سازمان‌ها این است که اطلاعات چگونه برای روسای سازمانی درباره افراد شاغل در زیردست‌شان ارائه و چگونه سیستم‌های پاداش ایجاد شود که با منافع کارفرمایان و کارگزاران همسو باشد. جست‌وجو برای اطلاعات منجر به معیارهای عملکرد می‌شود: اعداد استاندارد شده‌ای که به شکلی کارا به کارگزاران داده خواهد شد که نشان می‌دهد

عاملان‌شان تا چه اندازه توانسته‌اند اهداف آنها را به‌خوبی پیش برند. همسوسازی انگیزه‌ها به معنای دادن پاداش‌های مالی به کارکنان است که منعکس‌کننده سوددهی شرکت است: هنگامی که شرکت‌ها سود بیشتری داشته باشند، کارکنان نیز پاداش بیشتری خواهند داشت. ادبیات مدیریت حرفه‌ای، نتیجه‌گیری‌های خود را از نظریه کارفرما کارگزار گرفته است: اینکه مدیریت، موضوع تعیین اهداف مشخص و سپس ارزیابی و انگیزه‌سازی است. مدیریت از سویی به سیستم‌های اطلاعات و گزارش‌نویسی و از سوی دیگر به پاداش‌های ساختاریافته هوشمندانه متکی است.

### مدیریت عمومی جدید

با آغاز دهه ۱۹۸۰، این نوع تفکر از شرکت‌های سودآور به سازمان‌های دولتی و غیرانتفاعی مانند دانشگاه‌ها و بیمارستان‌ها بسط یافت. نارضایتی از هزینه‌ها، ناخشنودی از نتایج یا به بیان ساده‌تر، میل به پس‌انداز موجب انتقادهای و بحث‌هایی شد که بر اساس آن مشکل سازمان‌ها این بود که نیاز داشتند عملکردی «بیشتر شبیه به یک کسب‌وکار» داشته باشند. این طبل جنگ حامیان آنچه بود که تحت عنوان «مدیریت عمومی جدید» شناخته می‌شد. در وهله اول کارفرمایان کسانی بودند که هزینه آژانس‌ها و سازمان‌های غیرانتفاعی را پرداخت می‌کردند: که در فقره دولت، کارفرمایان همان مالیات‌دهندگان بودند. اکنون دانش‌آموزان و بیماران و ارباب‌رجوعان سازمان‌ها را به عنوان مشتریان خود در نظر می‌گرفتند.

یک دشواری برای کسانی که به دنبال ایجاد سازمان‌هایی شبیه به یک کسب‌وکار بودند این بود که هیچ سازوکار قیمتی وجود نداشت که تعیین کند آیا افرادی که وجوه مالی مورد نیازشان را تامین می‌کنند در ازای پول‌شان بهای خوبی دریافت خواهند کرد. در یک بازار رقابتی، مصرف‌کنندگان می‌توانند قیمت کالاها و خدمات را با توجه به کیفیت

محصولات ارائه شده مقایسه کنند و درباره اینکه چه چیزهایی بخرند، تصمیم‌های آگاهانه بگیرند. قیمت‌ها در بر دارنده اطلاعات بسیاری به صورت خلاصه و شفاف است اما مالیات‌دهندگان چگونه مدارس، دانشگاه‌ها، بیمارستان‌ها، سازمان‌های دولتی یا سازمان‌های خیریه را ارزیابی می‌کردند؟

برای حل این مشکلات، کسانی که به دنبال ایجاد سازمان‌های غیرانتفاعی شبیه به کسب‌وکار بودند، سه استراتژی را پیشنهاد دادند: اولین استراتژی تلاش برای توسعه شاخص‌هایی است که عملکرد را اندازه‌گیری و سنجش کند که می‌تواند جایگزینی برای قیمت نیز باشد. [۴] دوم، ارائه پاداش‌ها و تنبیهات مالی بر اساس عملکرد اندازه‌گیری شده به افراد شاغل در این سازمان‌هاست. سوم، ایجاد رقابت میان ارائه‌دهندگان است که شاخص‌های عملکرد آنها «شفاف» باشد، به معنای اینکه این شاخص‌ها در دسترس عموم قرار گیرد. به طور خلاصه، این ایده شرایطی شبیه بازار در دولت و بخش‌های غیرانتفاعی ایجاد و در نتیجه آنها را «بیشتر شبیه به یک کسب‌وکار» کرد. این شیوه تفکر «مدیریت عمومی جدید» نام دارد که بازتاب روند گسترده‌ای از وارد کردن اصول از اقتصاد خرد به مدیریت عمومی و سیاست‌گذاری عمومی است. [۵]

از همان آغاز، منتقدانی بودند که می‌کوشیدند تا توجهات را به سوی اشکالات گزاره‌های این رهیافت جلب کنند؛ اقتصاددانانی مانند بنیت هولمستروم<sup>۱</sup> و پُل میلگرام<sup>۲</sup> و همچنین هنری مینتزبرگ<sup>۳</sup> که استاد مدیریت دانشگاه مک‌گیل در مونترال بود از جمله این منتقدان بودند. [۶] مینتزبرگ در اواسط دهه ۱۹۹۰ خاطر نشان کرد که مفهوم مدیریت که توسط طرفداران مدیریت عمومی جدید مورد پذیرش قرار گرفته، در حقیقت یک کاریکاتور ساده شده است از آنچه که مدیران موثر در شرکت‌های بخش خصوصی واقعاً انجام می‌دهند. این تفکر با آنچه که

1. Bengt Holmström

2. Paul Milgrom

3. Henry Mintzberg

به بسیاری از دانشجویان مدیریت در مدارس کسب‌وکار و در ادبیات فراگیر مشاوره کسب‌وکار تدریس می‌شد، مطابقت داشت. او حتی در همان زمان نیز این بحث را مطرح کرد که این روند برای دولت و سازمان‌های غیرانتفاعی مناسب نیست. شرکت‌های تجاری دارای بخش‌هایی هستند که هر واحد دارای یک مأموریت واضح در ارائه مجموعه مشخصی از محصولات و خدمات است اما سازمان‌های دولتی و غیرانتفاعی با اهداف متعددی مشخص می‌شوند که تفکیک و اندازه‌گیری (سنجش) آنها دشوار است. برنامه‌های مدیریت عمومی جدید دارای راه‌حل‌های محتمل برای رویارویی با واحدهای دولتی است که تنها یک محصول یا خدمت مانند صدور گذرنامه ارائه می‌دهد؛ اما این امر بیشتر یک استثناست تا قاعده. علاوه بر این، در هر کسب‌وکاری، معیارهای مالی روشنی برای موفقیت و شکست وجود دارد: هزینه‌ها و مزایا را می‌توان برای تعیین میزان سود مقایسه کرد و مدیران احتمالاً می‌توانند بر این اساس پاداش دهند اما در سازمان‌های دولتی و غیرانتفاعی به‌ندرت اهداف واحدی وجود دارد و نمی‌توان به سرعت آنها را اندازه گرفت. مثلاً، مدارس ابتدایی اهداف خود را در زمینه تدریس خواندن و نوشتن و حساب ریاضی دارند و شاید بتوان این موارد را با آزمون‌های استاندارد ارزیابی کرد اما درباره اهداف مهم اما غیر قابل اندازه‌گیری مانند ترویج رفتار خوب، ایجاد حس کنجکاوی درباره جهان و تقویت فکر خلاق چطور؟ مشکل بزرگ‌تری نیز وجود دارد. شرکت‌ها در کسب‌وکارشان به دنبال سودآوری هستند و کارمندان‌شان در وهله اول برای کسب درآمد در مشاغل خود کار می‌کنند (این به آن معنا نیست که پول هدف غایی آنهاست، به عبارتی آنها در یک بخش بزرگ کار می‌کنند تا برای اهداف غیرمالی خود کسب درآمد کنند). افرادی که انتخاب می‌کنند تا برای سازمان‌های دولتی یا غیرانتفاعی مانند مدارس، دانشگاه‌ها، بیمارستان‌ها یا صلیب سرخ کار کنند نیز علاقه‌مندند تا از این طریق کسب درآمد کنند اما در عین حال تمایل دارند تا انگیزه و فعالیت بیشتری نسبت به تعهدات سازمانی که برای انجام آنها مأمور شده‌اند، داشته باشند: برای

تدریس، پژوهش، درمان و نجات. آنها به وسوسه پاداش‌های مالی به شکل‌های متفاوتی واکنش نشان می‌دهند زیرا انگیزه‌های آنها حداقل در میزانش متفاوت است. [۷]

### پاداش‌های بیرونی و درونی

بسیاری از مشکلات طرح‌های پرداخت مبتنی بر عملکرد را می‌توان در مفهوم بسیار ساده و کاملاً تحریف‌شده انگیزه انسان دنبال کرد؛ یعنی اینکه کسی تصور کند انگیزه افراد برای کار کردن تنها پاداش مالی است. برخی نسبت به پاداش‌های مالی بیرونی انگیزه کمتری دارند تا انواع مختلف پاداش‌های روحی-ذاتی شامل تعهدشان به اهداف سازمانی که در آن مشغول به کارند یا شیفتگی به پیچیدگی کاری که انجام می‌دهند که آن را چالش‌برانگیز، جالب و سرگرم‌کننده می‌کند. وجود انگیزه‌های ذاتی مانند انگیزه‌های بیرونی برای هر کسی که کارکنان را در وظایف پیچیده مدیریت می‌کند، آشکار است. این موضوع در اواسط دهه ۱۹۷۰ توسط روان‌شناسان بیان شد و از آن زمان تاکنون توسط اقتصاددانانی همچون ژان تیرول<sup>۱</sup>، برنده جایزه نوبل اقتصاد دوباره کشف و به رسمیت شناخته شد. [۸]

این ساده‌انگاری است که تصور شود افراد صرفاً با میل به پول بیشتر انگیزه می‌گیرند، همان‌طور که ساده‌لوحانه است که تصور شود افراد فقط با پاداش‌های ذاتی انگیزه پیدا می‌کنند. مساله، فهم این موضوع است که هر یک از این انگیزه‌ها تا چه اندازه موثرند و کدام یک بیشترین تاثیر را دارد؛ این موضوعی است که در سال‌های اخیر توسط دانشمندان علوم اجتماعی مورد توجه قرار گرفته است.

به طور کلی، پاداش‌های بیرونی-پرداخت مبتنی بر عملکرد و پرداخت تشویقی و پاداش‌ها- در سازمان‌های تجاری و با هدف اصلی کسب درآمد، بسیار تاثیرگذار است. همچنین این سازمان‌ها زمانی خوب کار می‌کنند که وظیفه‌ای که باید انجام شود، مجزا و به‌راحتی قابل

1. Jean Tirole

اندازه‌گیری باشد و همچنین مانند تولید برخی از کالاهای استاندارد در خط مونتاژ چندان اهمیت درونی نداشته باشد.

برخی پاداش‌ها انگیزه درونی را افزایش می‌دهند. مثلاً هنگامی که به جای اعمال کنترل از پاداش‌های کلامی در وهله اول برای انتقال اطلاعات استفاده می‌شود («کارت را خوب انجام دادی!») [۹] یا هنگامی که جوایزی پس از کسب دستاوردهای عالی داده می‌شود، بدون اینکه از قبل به عنوان یک انگیزه از آن صحبتی شده باشد. [۱۰] یا در شاخه‌های علوم و پژوهش، زمانی که جوایز و عناوین افتخاری به افراد برای به رسمیت شناختن دستاوردهای بلندمدتشان اعطا می‌شود. [۱۱] به طور گسترده‌تر، دستمزدهای بالاتر از سطح بازار اگر در حکم نشانه‌ای از قدردانی سازمان از عملکرد کارکنانش قلمداد شود، می‌تواند انگیزه درونی کارکنان را تقویت کند. [۱۲] انگیزه درونی و بیرونی در شرایطی می‌تواند موثر باشد که نتایج حاصل از پاداش‌ها با حس رسالت عامل همسو باشد: مثلاً هنگامی که بیمارستان‌ها برای ثبت سوابق ایمنی بهتر، پاداش دریافت می‌کنند.

اما هنگامی که سازمان‌های ماموریت‌محور تلاش می‌کنند که از پاداش‌های بیرونی به عنوان وعده‌هایی برای پرداخت مبتنی بر عملکرد استفاده کنند، نتیجه می‌تواند زیان‌آور باشد. استفاده از پاداش‌های بیرونی برای فعالیت‌های مبتنی بر علاقه درونی به این می‌انجامد که افراد به جای تمرکز بر منفعت درونی خود به وظایف یا ماموریت‌های بزرگ‌تری که بخشی از آن هستند، بر پاداش‌ها تمرکز کنند. نتیجه، «بیرون راندن» انگیزه درونی است: به آنها آموخته می‌شود که به وظایف شغلی خود در درجه اول به عنوان ابزارهایی برای رسیدن به اهداف پولی فکر کنند که در نتیجه آنها علاقه خود به انجام کار را به خاطر ماموریت بزرگ‌تر سازمان از دست می‌دهند. [۱۳] از سوی دیگر، آنها ممکن است پیشنهاد پرداخت در برابر عملکرد را توهینی به اخلاق حرفه‌ای خود قلمداد کنند زیرا این کار دلالت بر این امر دارد که آنها

کارشان را فقط به خاطر پول انجام می‌دهند. بنابراین تصور می‌شود که عملکرد پاداش‌های خارجی زمانی معنا دارد که فرد یک سرمایه‌گذار باشد، نه یک آموزگار یا پرستار. تلاش برای تبدیل همه‌چیز به یک کسب‌وکار مانع از این می‌شود که یک کسب‌وکار واقعی شکل بگیرد. در حقیقت، این امر مانع کسب‌وکارهای واقعی است. جالب اینجاست که حتی شرکت‌ها یکی پس از دیگری به دنبال توسعه طرح‌های انگیزشی پرداخت مبتنی بر عملکرد اندازه‌گیری شده برای مدیران و کارکنان ارشد خود بودند و چنین برنامه‌هایی برای سازمان‌های دولتی و غیرانتفاعی نیز بسیار جذابیت داشت، هرچند نظریه‌پردازان ارشد رفتار کارفرما-کارگزار در پایان قرن بیستم به دنبال یافتن نقاط ضعف چنین سیستم‌های بودند. رابرت گیبونز<sup>۱</sup> که استاد اقتصاد سازمانی در دانشگاه ام‌آی‌تی<sup>۲</sup> بود در سال ۱۹۹۸ خاطرنشان کرد که در حقیقت کارفرما (مثلاً صاحب شرکت) از انواع خروجی‌های عامل (کارمند) بهره‌مند می‌شود و بسیاری از این خروجی‌ها به طور محسوسی قابل مشاهده و با اعداد و ارقام قابل محاسبه نیستند. مثلاً سازمان‌ها به کارمندانی که مشغول مشاوره یا کارهای گروهی هستند وابسته‌اند و اغلب هم با آنچه که کارمندان انجام می‌دهند اختلاف نظر دارند؛ به‌ویژه اگر منافع آنها تنها به حداکثر رساندن عملکرد اندازه‌گیری شده خود برای جبران خسارت باشد. از این رو، میان سهم قابل اندازه‌گیری و واقعی و سهم نهایی عامل یک شکاف وجود دارد. در نتیجه، عملکرد اندازه‌گیری شده (مانند افزایش تقسیم مزایا یا افزایش قیمت سهام شرکت) ممکن است در واقع منجر به این شود که سازمان کمتر از آنچه که واقعاً نیاز دارد از طرف کارمندان خود دریافت کند. علاوه بر این، انحراف گریزناپذیر انگیزه‌ها از جست‌وجو برای استانداردهای ساده و کمی در راستای اندازه‌گیری و پاداش عملکرد ایجاد شده است. گیبونز نتیجه‌گیری کرد که در بهترین حالت، مدل‌های اقتصادی که طیف گسترده‌ای از انگیزه‌های روان‌شناختی را -به دلیل اینکه کارگزاران برای کار خود پاداش دریافت می‌کنند- نادیده می‌گیرند،

1. Robert Gibbons

2. MIT: Massachusetts Institute of Technology

مفهومی ناقص از انگیزه‌ها ارائه می‌دهند. در بدترین حالت، «شیوه‌های مدیریت مبتنی بر مدل‌های اقتصادی ممکن است واقعیت‌های غیراقتصادی همچون انگیزه‌های درونی یا روابط اجتماعی را تضعیف کنند (یا حتی از بین ببرند)» [۱۴]

در اواخر قرن بیستم، دانشجویان رفتار سازمانی مانند گیبونز توجه جدی خود را به مشکلات و خطرات جذابیت انگیزه‌های بیرونی معطوف کردند. هرچند در آن زمان، طرح‌های مبتنی بر مفاهیم ساده از انگیزه، پاداش بیرونی و مدیریت عمومی جدید به‌خوبی تثبیت و پذیرفته شده بود.

این شیوه‌های مدیریتی در بخش شرکتی خصوصی آغاز شد اما به سرعت گسترش یافت؛ بیشتر از همه در آنگلوِسِفِر (بریتانیای کبیر، ایالات متحده آمریکا، استرالیا و نیوزیلند). در راستای تلاش برای بهبود مدیریت و کارایی بخش عمومی، دولت محافظه‌کار مارگارت تاچِر<sup>۱</sup>، مجموعه‌های اداری را ایجاد کرد؛ بعضی از این مجموعه‌ها توسط تجار و بعضی دیگر توسط مشاوران مدیریت تحت عناوین مختلفی مانند واحد بهره‌وری، واحد مدیریت مالی، دفتر ملی حسابرسی و کمیسیون حسابرسی اداره می‌شدند. این شیوه از بریتانیا به استرالیا و نیوزیلند و دیگر کشورهای سازمان همکاری و توسعه اقتصادی بسط یافت و توسط آموزگاران مدیریت و مشاوران به منزله ابزارهای ترقی‌دانشگاهی و مدل‌های «بهترین شیوه (عمل)» به فراسوی مرزها برده شد. [۱۵]

1. Margaret Thatcher

## فصل ۶

## انتقادات فلسفی

فرهنگ معیارها، همان‌طور که دارای حامیان خود در هر دو طیف چپ و راست سیاسی است، منتقدانی نیز در هر دو سوی طیف ایدئولوژیک دارد. از منظر چپ مارکسیستی، با برخی توجیهات می‌توان شاهد ترویج کاهش مهارت بود که طبق آن تغییرات در سازمان تولید توسط افراد رده‌بالا دارای اثر کاهش ارزش مهارت‌ها و تجارب افراد رده‌پایین در سیستم تولید است.<sup>[۱]</sup> از همین منظر، کار<sup>۱</sup> هم محدودتر می‌شود و از آنجا که انتخاب نیز با تعریف محدود اهداف که توسط دیگران دیکته شده، از بین رفته است، بیشتر دچار بیگانگی می‌شود.

## توهم عقل‌گرا

همچنین کالبدشکافی‌های پرمایه‌ای درباره پاسخگویی به مثابه اندازه‌گیری، از جانب اندیشمندان محافظه‌کار و لیبرال کلاسیک مانند مایکل اوکشات<sup>۲</sup>، مایکل پولانی<sup>۳</sup> و فردریش هایک<sup>۴</sup> انجام شده است که تجزیه

1. Work

2. Michael Oakeshott

3. Michael Polanyi

4. Friedrich Hayek

و تحلیل‌های آنها اخیراً توسط جیمز سی. اسکات<sup>۱</sup>، انسان‌شناس دانشگاه ییل و از علاقه‌مندانِ آنارشیست، مجدداً کشف و بررسی شده است. همه آنها میان دو شکل دانش - یکی انتزاعی و دیگری فرمولی - تمایز قائلند که یکی عملی‌تر و ضمنی‌تر است. دانش عملی یا ضمنی محصول تجربه است: می‌توان آن را آموخت، اما به شکل فرمول‌های کلی قابل انتقال نیست. دانش انتزاعی درست برعکس است؛ یک موضوع تکنیکی است که فرض می‌شود به‌سادگی قاعده‌مند، منتقل و به کار گرفته می‌شود. در مثال مشهور اوکشات، نوعی دانش انتزاعی و دستورالعمل وجود دارد که در کتاب‌های آشپزی یافت می‌شود، اما در حقیقت فهم استفاده از چنین دانشی («زدن تخم‌مرغ» «هم‌زدن مخلوط») نیازمند دانش عملی مبتنی بر تجربه است که نمی‌توان آن را از کتاب‌ها آموخت. اوکشات از «عقل‌گرایان» انتقاد کرد زیرا آنها تصور می‌کردند انجام و اداره امور انسانی، موضوع به‌کارگیری فرمول‌ها و دستورالعمل‌های درست است. دانش فنی، مستعد فرمول‌سازی دقیق است که به آن ظاهر قطعیت می‌دهد. در مقابل، او نوشت:

«(این) یکی از ویژگی‌های دانش عملی است که مستلزم فرمول‌بندی به این شکل نیست. به بیان معمول، انجام امور به شیوه‌ای عرفی یا سنتی یا ساده‌تر در عمل است. همین موضوع به چنین دانشی ظاهری بی‌دقت و در نتیجه بدون قطعیت می‌دهد زیرا مساله مبتنی بر نظر و احتمال است تا حقیقت.»

عقل‌گرا به حاکمیت تکنیک باور دارد که مطابق آن تنها دانش معتبر، دانش فنی است زیرا تنها این دانش است که استاندارد قطعیت را دارد که دانش واقعی را مشخص می‌سازد. از نظر اوکشات، خطای عقل‌گرایی، درک نکردن آن از ضرورت دانش عملی و درک ویژگی‌های شرایط است.<sup>[۲]</sup>

## علم‌زدگی

فردریش هایک انتقادی را نسبت به آنچه او «تظاهر به دانش»<sup>۲</sup> می‌نامید،

1. James C. Scott

2. The Pretense of Knowledge

ارائه کرد. او در نوشته‌های خود در اواسط قرن بیستم تلاش‌های سوسیالیستی برای برنامه‌ریزی اقتصادی در مقیاس بزرگ را به دلیل «علم‌زدگی» شان مورد نکوهش قرار داد. او بر این باور بود که تلاش‌های آنها در حقیقت مهندسی حیات اقتصادی است؛ گویی برنامه‌ریزان در موقعیتی هستند که همه ورودی‌ها و خروجی‌های مربوط به زیستن در یک جامعه پیچیده را می‌دانند. به گفته او، مزیت بازار رقابتی به افراد اجازه می‌دهد که نه تنها از دانش خود در شرایط محلی بلکه از آن برای کشف کاربردهای جدید برای منابع موجود یا تصور محصولات و خدمات ناشناخته و غیرمنتظره جدید استفاده کنند. به طور خلاصه، برنامه‌ریزی نه تنها در ارزیابی اطلاعات مرتبط و حتی پراکنده شکست خورد، بلکه مانع کشف کارآفرینی و نحوه برطرف کردن نیازهای خاص و چگونگی تولید اهداف جدید هم شد. [۳]

از قضا، همان‌گونه که تعدادی از منتقدان معاصر مشاهده کرده‌اند، ثبات اهداف قابل اندازه‌گیری که برای ثبات معیارها بسیار حیاتی است - هرچند این امر اغلب توسط سیاست‌مداران و سیاست‌گذارانی انجام می‌شود که از پیش سرسپردگی خود را به سرمایه‌داری اعلام کرده‌اند - بسیاری از نقایص ذاتی سیستم شوروی را تکرار می‌کند. درست همان‌طور که برنامه‌ریزان بلوک شرق اهداف خروجی در تولید هر کارخانه را تعیین می‌کردند، بوروکرات‌ها نیز اهداف عملکرد قابل اندازه‌گیری را برای مدارس، بیمارستان‌ها، نیروهای پلیس و شرکت‌ها مشخص می‌کنند. درست همان‌طور که مدیران شوروی با تولید کالاهای بنجل که اهداف عددی تعیین‌شده توسط صاحبان خود را برآورده می‌کرد، به این روند پاسخ دادند، مدارس، نیروهای پلیس و صاحبان کسب‌وکار نیز روش‌هایی برای تحقق سهم خود با تولید کالاهای ارزان یافتند: فارغ‌التحصیلی دانش‌آموزان با حداقل مهارت‌ها، تقلیل سطح یک دزدی بزرگ به یک دستبرد کوچک یا باز کردن حساب‌های جعلی برای مشتریان بانک. [۴] بخش خوبی از نقد هایدک به علم‌زدگی نیز - که آن را درباره بسیاری از اقتصاددانان معاصر وارد می‌داند - مربوط به ایدئولوژی معیارهاست. با از

پیش تعیین کردن مجموعه‌ای از اهداف محدود و ظاهراً قابل اندازه‌گیری، ثبات معیار، گسترده اهداف واقعی یک کسب‌وکار یا سازمان را کوتاه و ناقص می‌کند. همچنین مانع بروز روحیه کارآفرینی در سازمان‌ها می‌شود زیرا ممکن است اهداف و مقاصد جدید ارزشمندی را دنبال کنند که بخشی از معیارها نباشد.

می‌توان بینش تعدادی از اندیشمندان را در این گفته خلاصه کرد: «حساب‌گری (برنامه‌ریزی) دشمن تخیل است.» همان‌طور که اشاره کردیم، کارآفرینی به چیزی بستگی دارد که فرانک نایت<sup>۱</sup> اقتصاددان از آن به منزله «ریسک غیرقابل اندازه‌گیری<sup>۲</sup>» نام می‌برد زیرا منافع بالقوه نوآوری چیزی نیست که بتوان آن را به‌دقت محاسبه کرد. یا در فرمول آلفی کوهن<sup>۳</sup> که برای مدت‌های طولانی منتقد پرداخت مبتنی بر عملکرد بود «معیارها مانع ریسک‌پذیری - که همراه گریزناپذیر کشف و خلاقیت است - می‌شوند؛ در این شرایط، احتمال کمی وجود دارد که این شانس را داشته باشیم که با امکانات و احتمالات بازی کنیم و به دنبال ایده‌هایی برویم که در نهایت برای‌شان هیچ پرداختی صورت نگیرد.» [۵]

همان‌گونه که جیمز اسکات<sup>۴</sup> خاطر نشان کرده است، ویژگی بارز دانش عملی و بومی این است که «این دانش برای مقابله با مشکل موجود، به همان اندازه که نیاز است - نه بیشتر و نه کمتر - اقتصادی و دقیق است.» [۶] در مقابل، میزان دقت عددی وعده‌شده توسط معیارها چه‌بسا خیلی بیشتر از نیاز شاغلان واقعی باشد و دستیابی به این دقت به صرف وقت و تلاشی نیاز دارد که ممکن است ارزش آن را نداشته باشد. بنابراین، تلاش برای دقت ممکن است بی‌فایده باشد و امکان دارد به همین دلیل مورد خشم و انزجار افرادی قرار بگیرد که زمان و نبوغ خود را وقف این کار کرده‌اند.

1. Frank Knight

2. Unmeasurable Risk

3. Alfie Kohn

4. James Scott

همان‌گونه که فیلسوف انگلیسی، آیزیا برلین<sup>۱</sup> در مقاله‌ای با عنوان «در باب داوری سیاسی» گفته است «برای خواست یا توصیه دقت مکانیکی، حتی در مبادی و اصول در یک زمینه نمی‌توان کور بود یا موجب گمراهی دیگران شد.» در واقع، آنچه که برلین درباره داوری (قضاوت) سیاسی می‌گوید کاربرد گسترده‌تری دارد: قضاوت (داوری) نوعی مهارت برای درک ویژگی‌های منحصر به فرد یک وضعیت است، شامل استعدادی برای سنتز است تا تحلیل، «ظرفیتی برای به دست دادن یک الگوی کلی از وضعیت انسانی، آن هم به شیوه‌ای که در آن همه‌چیز به هم متصل است.» [۷] حسی برای کل و خاص بودن؛ و این دقیقاً چیزی است که معیارهای عددی قادر به ارائه آن نیستند.

### نقد کدوری<sup>۲</sup> بر تاچر

در سال ۱۹۸۷ دولت محافظه‌کار مارگارت تاچر برنامه‌های گسترده‌ای برای تغییر بودجه عمومی آموزش عالی ارائه کرد. طرح‌ها به دنبال افزایش «شاخص‌های عملکرد» جدید و مبتنی بر شواهدی بود که طبق آنها وزیران و بوروکراسی‌های‌شان تصمیم به تخصیص بودجه به برخی دانشگاه‌های خاص گرفته بودند. مورخ برجسته و نظریه‌پرداز محافظه‌کار بریتانیایی، الی کدوری به عنوان یکی از مهم‌ترین منتقدان این برنامه ظاهر شد. او نوشت: «پس از دو دهه حمایت بی‌چون و چرا و بیش از حد از دولت، امروز ما شاهد یک نارضایتی مبهم اما قدرتمند و بیتاب با شیوه‌های کار دانشگاه‌ها هستیم ... یک تمایل بی‌نام‌ونشان برای برخی فرمول‌ها و دستورالعمل‌ها - شاید علم بیشتر، فناوری اطلاعات بیشتر، پرسشنامه‌های بیشتر و ارزیابی بیشتر - که به صورت علمی یا بهتر است بگوییم به صورت جادویی اثبات کنند که وقت خود را تلف نمی‌کنند و با انتقال‌دهندگان صنعت، ارتباطی نزدیک و پیوسته دارند.» [۸] او از این موضوع بسیار شگفت‌زده بود که «یک دولت محافظه‌کار باید با یک سیاست دانشگاهی بسیار متنوع و با ایده‌آل‌ها و اهداف از پیش تعیین‌شده آغاز کند» و نتیجه گرفت که «در راستای توضیح

1. Isaiah Berlin

2. Elie Kedourie

موارد غیر قابل توضیح، می‌توان چنین نتیجه گرفت که این سیاست، برآیند تصمیم‌های آگاهانه نیست بلکه یک واکنش غیرارادی ناخودآگاه نسبت به روحیه مقاومت‌ناپذیر زمان‌هاست.» [۹] کدوری اعلام کرد که تحت لوای شعار «کارایی»، کلاهبرداری بزرگی انجام شده است زیرا «کارایی یک ویژگی کلی و انتزاعی نیست و همیشه با هدف مورد نظر در ارتباط است. یک کسب‌وکار زمانی کاراست که بازگشت عوامل به کار گرفته‌شده در تولید در مقایسه با سایر موارد بیشتر باشد. اما دانشگاه یک کسب‌وکار نیست.» [۱۰] تحت لوای همین ادعا و بهانه که دانشگاه یک کسب‌وکار و دولت نماینده مشتریان خود است، کدوری مشاهده کرد که این وزیر آموزش و پرورش بود که بر اساس معیارهای جعلی تصمیم می‌گرفت که چه چیزی ارزش آموزشی دارد. [۱۱]

### پیش به سوی پاسخگویی

در دهه بعد، «پاسخگویی» و «اندازه‌گیری عملکرد» به اصطلاحاتی رایج در میان رهبران کسب‌وکار، سیاستمداران و سیاست‌گذاران در ایالات متحده آمریکا بدل شد. در سال ۱۹۹۳، رئیس‌جمهور بیل کلینتون، قانون عملکرد و دستاوردهای دولت<sup>۱</sup> را امضا کرد و همه سازمان‌ها ملزم شدند تا اهداف و ماموریت‌ها، برنامه‌های استراتژیک بلندمدت و اهداف عملکرد سالانه همراه با توضیحاتی درباره مقیاس‌های استفاده‌شده برای سنجش پیشرفت در این اهداف را اجرا کنند. این قانون ابتدا توسط قانون‌گذاران جمهوری‌خواه طرح شد و سپس توسط یک رئیس‌جمهور دموکرات به تصویب رسید و از این رو از حمایتی دوطرفه برخوردار بود. [۱۲] در سال ۲۰۰۴ و زمان ریاست‌جمهوری جورج دبلیو. بوش، دولت عالی فدرال نام اداره کل حسابداری<sup>۲</sup> را به دفتر پاسخگویی دولت<sup>۳</sup> تغییر داد.

با این امر ما وارد دوران خود می‌شویم و از تاریخ و نظریه عملکرد اندازه‌گیری شده به عمل معاصر ورود می‌کنیم.

1. Government Performance and Results Act

2. General Accounting Office

3. Government Accountability Office

بخش سوم

# اندازه‌گیری نادرست همه چیز؟ مطالعات موردی



## فصل ۷

## کالج‌ها و دانشگاه‌ها

برای اولین مطالعه موردی در حوزه آموزش عالی از نقطه آغازین پژوهش‌های خودم درباره ثبات معیارها آغاز می‌کنم. این پژوهش که در بر دارنده بخش بزرگی از اقتصاد ملی و یک نهاد مرکزی متشکل از کلیه انجمن‌های پیشرفته، کالج‌ها و دانشگاه‌هاست، بسیاری از اشکالات برجسته و پیامدهای ناخواسته در کنار برخی از مزایای عملکرد اندازه‌گیری‌شده را نشان می‌دهد.

**بالا بردن معیار: همه باید به کالج بروند**

همین که یک بار از ثبات اندازه‌گیری استفاده کنیم، به‌سادگی باور می‌کنیم که هر چه بیشتر از آن استفاده کنیم بهتر است. هم سازمان‌های دولتی و هم غیرانتفاعی، روز به روز آمریکایی‌های بیشتری را تشویق می‌کنند تا تحصیلات خود را پس از دبیرستان ادامه دهند. مثلاً طبق نظر وزارت آموزش ایالات متحده آمریکا، «در دنیای امروز کالج دیگر یک جای لوکس نیست که صرفاً تعداد محدودی از

آمریکایی‌ها از عهده تامین هزینه‌های آن برآیند، بلکه یک ضرورت اقتصادی و مدنی و شخصی برای همه آمریکایی‌هاست.»<sup>[۱]</sup>

در میان سازمان‌های غیرانتفاعی بسیار، سازمانی حامل پیام مشابهی در همین زمینه، بنیاد لومینا<sup>۱</sup> است. ماموریت این سازمان گسترش دستیابی به دوره‌های آموزشی پس از دوره متوسطه است با این هدف که ۶۰ درصد آمریکایی‌ها تا پایان سال ۲۰۲۵ مدرک کالج، گواهینامه یا اعتبارنامه عالی تحصیلی پس از دوره متوسطه داشته باشند. این برنامه چنانکه بنیاد در وبسایت خود اعلام کرده است، تحت عنوان «کشور قوی‌تر»<sup>۲</sup> اجرا می‌شود:

«همه چیز درباره شواهد مربوط به یادگیری است - کمیت آن، پیگیری آن و تعیین محل‌هایی که در آنها یادگیری اتفاق می‌افتد یا نه ... همچنین لومینا با رهبران سیاسی ایالتی سراسر کشور در راستای تعیین اهداف و توسعه و اجرای برنامه‌های قوی دولتی برای دستیابی به اهداف تعیین شده همکاری می‌کند. تاکنون، ۲۶ ایالت، اهدافی دقیق و چالشی را مشخص کرده‌اند - از این تعداد ۱۵ ایالت در سال گذشته به این برنامه پیوسته‌اند. بیشتر این ایالت‌ها، گام‌های محکمی در این مسیر برای رسیدن به اهداف‌شان برداشته‌اند؛ از جمله سرمایه‌گذاری‌های خروجی‌محور، ارتقای آموزش مرتبط با توسعه و ارزان‌تر کردن آموزش عالی.»<sup>[۲]</sup>

بنیاد لومینا کاملاً تحت نفوذ معیارهاست و به نمایندگی از آنها عمل می‌کند؛ وبسایت این بنیاد اعلام می‌کند: «بنیاد لومینا به عنوان سازمانی که بر نتایج تمرکز دارد از معیارهای کشوری برای هدایت کار، اندازه‌گیری تاثیرات و نظارت بر پیشرفت کشور و حرکت به سوی اهداف ۲۰۲۵ استفاده می‌کند.»

ماموریت بنیاد لومینا با یک باور رایج و مشترک درباره نقش آموزش عالی در جامعه آمریکا همراه است: این باور که افراد بیشتری باید به کالج بروند، زیرا با این کار نه تنها درآمد شخصی افراد بیشتر می‌شود بلکه موجب رشد اقتصاد ملی خواهد شد.

1. Lumina Foundation

2. Stronger Nation

### افزایش تعداد برنده‌ها از ارزش برنده شدن می‌کاهد

آن ایمان و اهداف عملکرد ناشی از آن ممکن است به‌سادگی اشتباه گرفته شود. همان‌گونه که آلیسون وولف<sup>۱</sup>، اقتصاددان آموزشی در دانشگاه لندن اشاره می‌کند، افراد دارای مدرک دانشگاهی کارشناسی تمایل دارند به طور متوسط بیشتر از افراد بدون مدرک، کسب درآمد کنند. بنابراین در سطح فردی، تلاش برای گرفتن مدرک کارشناسی چه‌بسا موجب هوش اقتصادی شود. اما در سطح ملی، این باور که تحصیلات دانشگاهی بیشتر موجب بهره‌وری بیشتر می‌شود، یک تصور اشتباه است.<sup>[۳]</sup>

یکی از دلایل این امر آن است که تحصیلات تا اندازه زیادی یک کالای موقعیتی است؛ دست‌کم در هنگام ورود به بازار کار. در مورد کارفرمایان بالقوه، مدارک دانشگاهی همچون نشانه‌ها عمل می‌کنند: آنها مانند راه میان‌برند و به کارفرمایان اجازه می‌دهند تا متقاضیان اولیه برای شغل را با توجه به مدارک آنها رتبه‌بندی کنند. تمام کردن دوره دبیرستان نشان‌دهنده سطح متوسطی از صلاحیت فکری و برخی ویژگی‌های شخصیتی همچون پشتکار است. به پایان رساندن کالج تا حدی نشان‌دهنده این ویژگی‌ها در سطحی بالاتر است. در جامعه‌ای که تنها اقلیت کوچکی توانسته‌اند کالج را به پایان برسانند، داشتن مدرک کارشناسی، معیاری قطعی از برتری به شمار می‌رود. اما هر چه درصد افراد دارای مدرک کارشناسی بیشتر شود، ارزش آن به عنوان شیوه و ابزار دسته‌بندی کاهش می‌یابد. در عوض آنچه‌گه اتفاق می‌افتد این است: مشاغلی که زمانی به دیپلم نیاز داشتند، اکنون به مدرک کارشناسی نیاز دارند. این موضوع به این دلیل نیست که مشاغل به لحاظ شناختی دشوار شده‌اند یا نیازمند مهارت‌های بیشتر هستند بلکه به این دلیل است که کارفرمایان می‌توانند از میان تعداد بسیار زیادی متقاضی شغل دارای مدرک کارشناسی دست به انتخاب بزنند و دیگران را حذف کنند. نتیجه این است که هم دستمزد کسانی که دارای مدرک کالج نیستند کاهش می‌یابد و هم فارغ‌التحصیلان کالج را

1. Alison Wolf

در مشاغلی قرار می‌دهد که عملاً از تحصیلات و آموخته‌های تحصیلی خود بهره‌ای نمی‌برند. [۴] این امر، منجر به یک مسابقه تسلیحاتی موقعیتی می‌شود: به محض اینکه این سخن پخش شود که مدرک کالج، کلید ورود به مشاغل حتی متوسط است، افراد بیشتر و بیشتر به دنبال مدرک می‌روند.

بنابراین، انگیزه‌های شخصی برای افزایش تعداد افرادی که برای کسب مدرک تلاش می‌کنند وجود دارد. در همین راستا، سازمان‌های دولتی و خصوصی عملکرد را با هدف افزایش حضور در کالج و فارغ‌التحصیلی در آن اندازه‌گیری می‌کنند.

### معیارهای بالاتر به واسطه استانداردهای پایین‌تر

اما واقعیت این است که بیشتر آمریکایی‌هایی که وارد کالج می‌شوند لزوماً آماده این کار نیستند، یا همه آمریکایی‌ها قادر به گرفتن یک مدرک قابل توجه نیستند.

در واقع، هیچ نشانه‌ای مبنی بر اینکه دانش‌آموزانی که دبیرستان را ترک می‌کنند آماده تحصیل در سطح کالج هستند وجود ندارد. [۵] یکی از معیارهای آمادگی برای ورود به کالج، عملکرد دانش‌آموزان در آزمون‌های پیشرفت تحصیلی همچون اس‌ای‌تی<sup>۱</sup> و ای‌سی‌تی<sup>۲</sup> است که احتمال موفقیت آنها را در کالج پیش‌بینی می‌کند (آنها تا حدودی آزمون‌های استعداد تحصیلی هستند). در بیشتر موارد، این آزمون‌ها از دانش‌آموزان دبیرستانی گرفته می‌شود که امید به ادامه تحصیل در سطوح بالاتر دارند، اگرچه در راستای تلاش برای افزایش موفقیت دانش‌آموزان، برخی ایالت‌ها ملزمند تا از دانش‌آموزان بیشتری این امتحانات را بگیرند؛ احتمالاً یک مورد از رابطه علت و معلولی نادرست است، اما دانش‌آموزانی که در این آزمون‌ها شرکت کرده‌اند به پیشرفت

1. SAT: Scholastic Assessment Test

آزمون استاندارد است جهت ورود به کالج و بعضی از دانشگاه‌های آمریکا.

2. ACT: American College Test

یک آزمون استاندارد در آمریکاست که آمادگی دانش‌آموزان را برای ورود به کالج ارزیابی می‌کند.

بیشتر تمایل داشته‌اند. از این رو، به اشتباه این‌طور استدلال می‌شد که هر چه تعداد دانش‌آموزان بیشتری در این آزمون‌ها شرکت کنند، میزان موفقیت نیز بیشتر خواهد بود. ایراد اینجاست که احتمال اینکه دانش‌آموزان در مرحله ابتدایی یا آغاز آزمون‌ها عملکرد بهتری داشته باشند، زیاد است. به همین دلیل است که سیاست‌گذاران درباره علت معلول اشتباه می‌کنند. آزمون ای‌سی‌تی از چهار بخش موضوعی آزمون می‌گیرد: انگلیسی، ریاضی، خواندن و علوم. نهادی که آزمون ای‌سی‌تی را برگزار می‌کند، معیارهای نمرات را به‌گونه‌ای توسعه داده است که نشان دهد فردی که آزمون می‌دهد «آمادگی جدی برای کار در دوره کالج را دارد.» از میان کسانی که اخیراً در این آزمون شرکت کرده‌اند، یک سوم در هیچ یک از چهار حوزه، نمرات معیار را کسب نکرده‌اند و تنها ۳۸ درصد حداقل در سه حوزه از چهار حوزه توانسته‌اند به نمرات معیار دست یابند. به‌طور خلاصه، می‌توان گفت بیشتر افرادی که تمایل به ادامه تحصیل داشته‌اند، نشان دادند که توانایی انجام این کار را ندارند. [۶]

نتایج قابل پیش‌بینی هستند؛ اگرچه کمتر کسی آنها را می‌پذیرد. بیشتر دانش‌آموزانی که وارد مجموعه کالج یا دوره‌های چهار ساله کالج می‌شوند، به‌طور کامل آماده نیستند و نیاز به دوره‌های جبرانی دارند. این دوره‌ها (که امروزه با حسن تعبیر با عنوان «تکمیلی» شناخته می‌شوند) را دانش‌آموزان باید در دبیرستان بیاموزند. یک سوم دانش‌آموزانی که وارد دوره‌های کالج شده‌اند، در کلاس‌های تکمیلی خواندن و ۵۹ درصد در کلاس‌های تکمیلی ریاضی شرکت کرده‌اند. [۷]

دانش‌آموزانی که آمادگی کافی برای ورود به دوره‌های کالج را کسب نکرده‌اند از نهادهایی که به آنها وارد شده‌اند درخواست دیگری نیز دارند و به این ترتیب هزینه‌های آموزش کالج افزایش می‌یابد: افزایش تعداد پردیس‌ها در مراکز «عالی آموزشی» یک پاسخ غیرمستقیم به نیاز برای برنامه‌های کمک‌درسی بیشتر در مهارت نوشتاری و دیگر مهارت‌هایی است که دانش‌آموزان برای کار در سطح دانشگاه به اندازه

کافی از آنها برخوردار نشده‌اند.

کالج‌ها اعم از دولتی و خصوصی تا اندازه‌ای بر اساس نرخ فارغ‌التحصیلی سنجش و پاداش داده می‌شوند و این یکی از معیارهای رتبه‌بندی کالج‌ها و در برخی موارد پاداش به آنهاست (تشویق بنیاد لومینا از طرف مقامات ایالتی به دلیل فعالیت در برنامه «تامین مالی خروجی محور» را به یاد آورید). آنچه اتفاق می‌افتد این است که خروجی‌ها به دنبال تامین مالی (بودجه) می‌آیند. کالج با اجازه دادن به تعداد بیشتری دانش‌آموز برای گذراندن دوره‌ها و امتحانات، به روشنی پاسخگویی خود را به واسطه معیار عملکرد عالی خود نشان می‌دهد. آنچه چندان شفاف نیست، پایین آمدن استانداردهای مورد نیاز برای فارغ‌التحصیلی است. [۸] دوره‌های بیشتر با نیازهایی که به‌سادگی برآورده می‌شود، ارائه خواهد شد. روی استادان فشارهایی وارد می‌شود؛ گاه آشکار و گاه ضمنی [۹]، که در نمره دادن سخاوتمند باشند. بخش بزرگی از اعضای هیات علمی دانشگاه‌ها نیز استادان حق‌التدریس هستند و اگر بخش بزرگی از دانشجویان کلاس‌های این استادان در امتحانات ناکام بمانند (حتی اگر عملکرد استادان خوب باشد)، احتمال اینکه قراردادشان تمدید شود، بسیار کم است.

پس دانش‌آموزان بیشتری وارد کالج‌ها و دانشگاه‌ها می‌شوند. نتیجه این است که دانش‌آموزان بسیاری بدون اینکه توانایی کار در سطوح کالج و دانشگاه را داشته باشند به این مراکز وارد می‌شوند و برخی نیز در این نهادها ثبت‌نام می‌کنند اما تحصیلات خود را تا گرفتن مدرک به پایان نمی‌رسانند؛ پدیده‌ای رایج و روبه‌رشد که هزینه‌های قابل توجهی مانند شهریه و هزینه‌های زندگی برای دانشجوی فراهم می‌آورد و موجب از بین رفتن درآمد نیز می‌شود. [۱۰] به نظر می‌رسد نرخ بالای ترک تحصیل‌کنندگان نشان‌دهنده این باشد که تعداد زیادی از دانش‌آموزان به کالج وارد می‌شوند [۱۱] و کسانی که مدرک تحصیلی کارشناسی دریافت می‌کنند موجب کاهش ارزش اقتصادی می‌شوند زیرا نشانه‌های کمتر و کمتری به کارفرمایان بالقوه درباره توانایی و دستاورد خودشان

می‌دهند. [۱۲] با درک این موضوع، دانش‌آموزان آینده‌نگر و والدین‌شان به دنبال پذیرش از هر کالژی نیستند بلکه به دنبال کالجهایی هستند که در رتبه‌بندی جایگاه بالاتری دارند [۱۳] و این امر به نوبه خود منجر به مسابقه تسلیحاتی در رتبه‌بندی کالج‌ها می‌شود که به این موضوع بازخواهیم گشت.

کاهش استانداردها برای کسب مدرک کارشناسی به این معناست که درصد افرادی که مدرک کالج را گرفته‌اند به عنوان یک شاخص «سرمایه انسانی» قلمداد و به واحد گمراه‌کننده اندازه‌گیری برای تحلیل سیاست عمومی بدل شود. اقتصاددانان فقط می‌توانند آنچه را که اندازه می‌گیرند ارزیابی کنند و لازم است آنچه را که اندازه‌گیری می‌کنند استاندارد باشد. بنابراین اقتصاددانانی که روی «سرمایه انسانی» و سهم آن در رشد اقتصادی کار می‌کنند (و تقریباً همیشه نتیجه می‌گیرند که اقتصاد به فارغ‌التحصیلان بیشتری نیاز دارد)، اغلب از نرخ فارغ‌التحصیلان کالج در حکم معیاری در دستیابی به «سرمایه انسانی» بهره می‌برند؛ غافل از اینکه همه مدارک کارشناسی یکسان نیستند و ممکن است بعضی از آنها بازتاب توانایی یا موفقیت نباشد. این امر، منجر به وجود قطعی عدم واقعیت می‌شود که برای رفع آن اقتصاددانان بدون سروصدا و با تمایلی نه چندان زیاد، دست به ترکیب معیارهای دشوار معتبر آماری با واحدهای اندازه‌گیری معتبر می‌زنند.

در همین راستا یک فرض این است که این تلاش برای افزایش ثبت‌نام در کالج‌ها و ادامه تحصیل و گرفتن مدرک که منجر به افزایش میانگین پیشرفت تحصیلی می‌شود، در نهایت به سطوح بالاتر رشد اقتصاد ملی خواهد انجامید. اما برخی اقتصاددانان برجسته در هر دو سوی اقیانوس اطلس - آلیسون وولف<sup>۱</sup> در انگلستان و دارون عجم‌اگلو<sup>۲</sup> و دیوید آتور<sup>۳</sup> در ایالات متحده آمریکا - به این نتیجه رسیده‌اند که چنین ادعایی دیگر درست نیست؛ البته اگر تاکنون درست بوده باشد. در عصری که فناوری

1. AlisonWolf

2. DaronAcemoglu

3. DavidAutor

جایگزین وظایفی شده که پیشتر توسط سرمایه انسانی در سطوح پایین و متوسط انجام می‌شد، رشد اقتصاد ملی بیشتر مبتنی بر نوآوری و پیشرفت فناوری است و نه میانگین پیشرفت تحصیلی به عنوان دستاورد افرادی که در صدر دانش، توانایی و مهارت قرار دارند. [۱۴] در دهه‌های اخیر، درصد جمعیت دارای مدرک کالج افزایش یافته اما نرخ رشد اقتصادی کاهش یافته است و اگرچه شکاف میان درآمد افراد دارای مدرک دانشگاهی و افراد بدون مدرک همچنان چشمگیر است اما کاهش میزان درآمد فارغ‌التحصیلان نشان می‌دهد که در حال حاضر اقتصاد با عرضه بیش از حد فارغ‌التحصیلان مواجه است. [۱۵] در عوض، کمبود کارگران در مشاغل پیشه‌وری مانند لوله‌کش، نجار و برق‌کار - مشاغلی که آموزش آنها از طریق کارآموزی و نه تحصیلات دانشگاهی انجام می‌شود - موجب شده است تا درآمد آنها از افراد دارای مدرک بیشتر باشد. [۱۶] مطمئناً، سیاست‌گذاری عمومی باید بیش از رشد اقتصادی و اثرات تحصیل در کالج در کسب ظرفیت بیش از خود آن مورد توجه قرار گیرد و این چیزی است که ما به دنبال کشف آن هستیم. در حال حاضر هدف معیار در افزایش تعداد فارغ‌التحصیلان حتی با معیارهای اقتصادی که آموزش عالی با آن سنجیده می‌شود، با تردید همراه است.

### فشار برای اندازه‌گیری عملکرد کالج‌ها

در دهه‌های اخیر، از زمانی که الی کدوری از سیاست متمرکزسازی دولت محافظه‌کار مارگارت تاچر انتقاد کرد، کنترل دولت مرکزی بر موسسه‌های آموزش عالی انگلیس گسترش و شدت یافت. بخش اعظم این کنترل به شکل مدیریت از طریق معیارهای عملکرد و نتایج در بسیاری از حوزه‌های آموزش نیز مخرب بود.

در انگلستان نیز مانند هر جای دیگری، بخش بزرگی از جمعیت در همراهی با اهداف دولت به دنبال تحصیل در دانشگاه هستند. در سال ۱۹۷۰ کمتر از ۱۰ درصد مردان و زنان در هر گروه سنی وارد دانشگاه می‌شدند. تا سال ۱۹۹۷ این میزان به یک سوم رسید و تا سال ۲۰۱۲

حدود ۳۸ درصد افراد ۱۹ ساله به نوعی در آموزش عالی ثبت‌نام کرده بودند. [۱۷] پرداخت هزینه برای آنها کاری طاقت‌فرساست و در سال‌های اخیر افزایش هزینه‌ها در قالب شهریه به دانشجویان (یا خانواده‌ی‌شان) فشار زیادی وارد کرده است. اما هزینه‌های دولت همچنان قابل توجه است و برای کاهش هزینه‌ها و دستیابی به «ارزش» در قالب کنترل روزافزون برای پرداخت در راستای نتایج تعیین‌شده تلاش می‌کند. این عملکرد از طریق معیارهایی ارزیابی می‌شود که بر خروجی اندازه‌گیری‌شده در هر بخش و هر نهاد تمرکز دارند.

در تلاش برای کسب «ارزش»، دولت‌های پیاپی انگلیس، مجموعه‌ای از سازمان‌های دولتی را برای ارزیابی دانشگاه‌های کشور با عنوان «سازمان تضمین کیفیت<sup>۱</sup>» ایجاد کرده‌اند. [۱۸] بررسی‌هایی از کیفیت تدریس مانند «ارزیابی کیفیت تدریس» انجام شده است؛ این ارزیابی به طور گسترده با رویه‌هایی دنبال‌شده و تکالیف بایگانی‌شده صورت گرفته است که ارتباط چندانی هم با تدریس واقعی ندارد. [۱۹] یک نتیجه روشن این است که استادان مجبورند زمانی را که باید صرف تدریس و پژوهش کنند به تکالیف و کاغذبازی اختصاص می‌دهند و روزبه‌روز به تعداد کارمندانی که در این مشاغل مشغول کارند افزوده می‌شود؛ پست تازه‌ای تحت عنوان «ماموران تضمین کیفیت» ایجاد شده است و افراد در آن با عنوان «ارزیابی کیفی پژوهش» به گردآوری و تحلیل داده‌ها مشغولند که بعدها به چارچوب عالی پژوهش تغییر نام داد. [۲۰] هزینه این ارزیابی‌ها در سال ۲۰۰۲ تنها در انگلستان حدود ۲۵۰ هزار پوند تخمین زده شده است. [۲۱] رشد قارچ‌گونه کارمندان اداری در کشورهای دیگری مانند استرالیا که سیستم مشابه اندازه‌گیری عملکرد را پذیرفته بودند نیز اتفاق افتاد. در بیشتر این سیستم‌ها، معیارها، زمان و انرژی افراد مشغول تدریس و پژوهش را گرفته و آنها را به سوی گردآوری و انتشار داده‌ها برای ارزیابی پژوهش سوق داده و به جای اینکه زمان و انرژی‌شان را صرف امور حقیقی کنند، آنها را به سوی مستندسازی منحرف ساخته‌اند. [۲۲] جست‌وجو برای داده‌های بیشتر به معنای داده‌های بیشتر،

1. Quality Assurance Agency

مدیران بیشتر، بوروکراسی بیشتر و سیستم‌های نرم‌افزاری بیشتر و گران‌تر است. جالب اینکه تمام این موارد به نام کنترل هزینه انجام می‌شود اما خود به افزایش هزینه می‌انجامد. همتای این سازمان‌ها در ایالات متحده آمریکا، سازمان‌های معتبری هستند که به کالج‌ها و دانشگاه‌های آمریکایی مشروعیت می‌بخشند. هرچند کالج‌ها و دانشگاه‌ها در حوزه منطقه‌ای هستند اما از آنجا که بودجه فدرال را دریافت می‌کنند نیازمند اعتباربخشی توسط چنین سازمان‌هایی هستند، آنها همچنین به عنوان ابزارهای دولت فدرال نیز خدمت می‌کنند. [۲۳] هرچند آنها بودجه را به شیوه همتایان بریتانیایی خود کنترل نمی‌کنند اما به هر حال نقش عمده‌ای در این امر دارند. در دهه‌های اخیر، این نقش عبارت است از تحت فشار قرار دادن کالج‌ها و دانشگاه‌هایی که به آنها اعتبار داده‌اند و حالا آنها موظفند اندازه‌گیری‌های دقیق‌تری از عملکردشان تحت عنوان «ارزیابی» داشته باشند. [۲۴]

پاداش برای عملکرد اندازه‌گیری‌شده در آموزش عالی با حمایت افرادی که به دنبال این بودند که دانشگاه‌ها «همچون یک کسب‌وکار» شوند، تقویت شد اما کسب‌وکارها دارای محدودیت‌های خاص زمانی و پولی برای اندازه‌گیری هستند که در برخی مواقع منجر به کاهش سود می‌شود. از آنجا که دانشگاه‌ها و نهادهای غیرانتفاعی دارای چنین سودی نیستند، سازمان‌های دولتی یا اعتباربخشی یا ریاست اداری دانشگاه‌ها می‌توانند معیارها را بدون وقفه گسترش دهند. [۲۵] اثر این کار افزایش هزینه‌ها و یا انتقال آنها از کارفرمایان به مدیران است که معمولاً به مورد دوم (مدیران) منتقل می‌شود. در دهه‌های اخیر یافتن دانشگاهی که نسبت مدیران به استادان و مدیران به دانشجویان به طور نجومی افزایش نیافته باشد، دشوار است. [۲۶] همین امر در سطح ملی نیز صادق است.

### مسابقه تسلیحاتی رتبه‌بندی

یکی دیگر از مجموعه‌هایی که به طور فزاینده‌ای متاثر از عملکرد معیارها در حوزه آموزش عالی است، رتبه‌بندی دانشگاه‌هاست. رتبه‌بندی‌ها شکل‌های مختلفی به خود می‌گیرند؛ در سطح بین‌المللی، رتبه‌بندی

شانگهای ژیاوتونگ<sup>۱</sup> برای «رتبه‌بندی علمی دانشگاه‌های جهان» وجود دارد (این رتبه‌بندی در راستای ارائه یک «معیار جهانی» توسط دولت چین تهیه شده تا دانشگاه‌های چین بتوانند پیشرفت خود را در دستیابی به «پژوهش علمی دشوار» ارزیابی کنند و از این رو، این رتبه‌بندی ۹۰ درصد وزن خود را به نشریات و جوایز در علوم طبیعی و ریاضیات می‌دهد) [۲۷] و مجله تایمز هایلر اجوکیشن ساپلیمنت<sup>۲</sup> تلاش می‌کند دانشگاه‌های جهان را بر اساس تدریس، پژوهش (شامل مجلدات، نشریات و ارجاعات) و «چشم‌انداز بین‌المللی» رتبه‌بندی کند. در ایالات متحده آمریکا مهم‌ترین رتبه‌بندی‌ها متعلق به اخبار آمریکا و گزارش جهانی<sup>۳</sup> است که با رقابت میان فوربس<sup>۴</sup>، نیوزویک<sup>۵</sup>، پرینستون ریویو<sup>۶</sup> و کیپلینگر<sup>۷</sup> (که تلاش می‌کند میان کیفیت و قیمت تعادل ایجاد کند) و تعدادی دیگر معنا پیدا می‌کند. این رتبه‌بندی‌ها -یا آنچنان که در بریتانیا با عنوان «جداول رده‌بندی» شناخته می‌شود- منبع مهمی از اعتبار به شمار می‌رود: فارغ‌التحصیلان و اعضای هیات امناء، همچون حامیان و دانشجویان، مشتاقند تا موسسه آنها رتبه بالایی در این رتبه‌بندی‌ها کسب کند. حفظ و بهبود رتبه‌بندی موسسه‌ها بدل به یک اولویت برای روسای دانشگاه و مدیران ارشد آنها می‌شود. [۲۸] در حقیقت با برخی از روسای دانشگاه‌های آمریکا قراردادهایی بسته می‌شود که طبق آن اگر روسا بتوانند رتبه دانشگاه را بالا ببرند به آنها پاداش می‌دهند. درباره سایر مدیران ارشد نیز باید گفت از آنجا که یکی از عوامل موثر بر رتبه‌بندی، نمرات کسب‌شده دانشجویان ورودی است، مسئول پذیرش بر اساس بخشی از نمرات دانشجویان پذیرفته‌شده پاداش می‌گیرد. [۲۹]

1. Shanghai Jiao Tong

2. Times Higher Education Supplement

«تایمز هایلر اجوکیشن ساپلیمنت» یک نشریه انگلیسی است که همه ساله دانشگاه‌های جهان را رتبه‌بندی و دانشگاه‌های برتر جهان را معرفی می‌کند. م

3. US News and World Report (USNWR)

4. Forbes

5. Newsweek

6. Princeton Review

7. Kiplinger

اخیرا من متوجه این موضوع عجیب شده‌ام که یک دانشگاه متوسط آمریکایی در تمام شماره‌های مجله «تاریخچه آموزش عالی»<sup>۱</sup>، به انتشار مطالب مهمی که اعضای هیات علمی روی آنها کار می‌کنند، می‌پردازد. از آنجا که این مجله بیشتر توسط دانشگاهیان و مدیران دانشگاهی خوانده می‌شود، من از هزینه‌های گزافی که این دانشگاه نه چندان ثروتمند صرف این تبلیغات اضافی و سطحی می‌کرد بسیار شگفت‌زده شدم. سپس این موضوع به شدت مرا تحت تاثیر قرار داد: رتبه‌بندی اخبار آمریکا و گزارش جهانی تا اندازه زیادی بر مبنای نظرسنجی از روسای کالج‌ها درباره رتبه‌بندی و اعتبار دیگر دانشگاه‌هاست. این معیار از اعتباری نامطمئن برخوردار است زیرا بیشتر روسا از تحولات موسسه‌های دیگر بی‌اطلاع هستند. هدف این کمپین تبلیغاتی بالا بردن سطح آگاهی در دانشگاه‌ها در تلاش برای تقویت اعتبار رتبه‌بندی‌های اخبار آمریکا و گزارش جهانی است.

همچنین دانشگاه‌ها هزینه‌های بسیاری صرف بروشورهای گران‌قیمت برای ارائه دستاوردهای سازمانی و علمی خود می‌کنند. این بروشورها برای روسای دانشگاه‌های دیگر که در نظرسنجی «یو.اس.ان.دابلو.آر» شرکت کرده‌اند، ارسال می‌شود. اگرچه دانشگاه‌ها - و برخی دانشکده‌ها مانند دانشکده حقوق - میلیون‌ها دلار آشکارنشدی صرف این نشریات بازاریابی می‌کنند اما شواهدی مبتنی بر اینکه این کار واقعا موثر است وجود ندارد. در حقیقت بیشتر این نشریات به صورت بازنشده توسط گیرندگان آنها به درون سطل زباله پرتاب می‌شود. [۳۰]

علاوه بر هزینه‌هایی که هیچ تاثیری در افزایش کیفیت تدریس یا پژوهش ندارد، افزایش اهمیت رتبه‌بندی منجر به ایجاد انواع جدیدی از بازی‌ها از طریق کارهای سطحی و ارتقای اعداد به واسطه حذف و تحریف داده‌ها شده است. یک پژوهش علمی جدید درباره دانشکده‌های حقوق در آمریکا چند نمونه از این دست ارائه می‌کند. «یو.اس.ان.دابلو.آر» دانشکده‌های حقوق را بر اساس نمرات آزمون ورودی دانشکده‌ها و

1. The Chronicle of Higher Education

میانگین نمرات دانشجویان تمام‌وقت پذیرفته‌شده رتبه‌بندی می‌کند. برای بهبود آمارها، دانشجویان با نمرات پایین‌تر یا به صورت «پاره‌وقت» یا به صورت «مشروط» پذیرش می‌شوند و به این ترتیب نمرات آنها در میانگین و رتبه‌بندی لحاظ نمی‌شود. از آنجا که نمرات دانشجویان انتقالی نیز محاسبه نمی‌شود، بسیاری از اداره‌های پذیرش دانشکده‌های حقوق خواهان پذیرش دانشجویانی از دیگر دانشکده‌ها با رتبه‌بندی و نمرات پایین‌تر هستند تا پس از سال اول به آنجا منتقل شوند. نسبت پایین دانشجویان به عضو هیات علمی نیز در رتبه دانشکده نقش موثری دارد اما از آنجا که این نسبت‌ها نیز به طور معمول در ترم پاییز محاسبه می‌شوند، دانشکده‌های حقوق اعضای هیات علمی را تشویق می‌کنند تا در ترم بهار مرخصی بگیرند. [۳۱] اینها تکنیک‌هایی برای بازی با سیستم رتبه‌بندی است که به هیچ وجه محدود به دانشکده‌های حقوق نیست و در بیشتر کالج‌ها و دانشگاه‌ها وضعیت به همین منوال است. [۳۲]

آیا ارزشش را دارد؟ برخی پژوهش‌های جدید نشان می‌دهد که اختلاف‌های اندک در رتبه‌بندی کالج‌ها، برخلاف باور مدیران این موسسه‌ها تاثیر اندکی در ثبت‌نام دانش‌آموزان دارد و اینکه منابع صرف‌شده برای بهبود رتبه‌بندی تناسبی با تاثیر واقعی آنها ندارد. [۳۳] اگر چنین است، پس چنین پیامی هنوز به گوش بسیاری از مسئولان دانشگاه نرسیده است.

### سنجش بهره‌وری دانشگاهی

در تلاش برای جایگزینی داوری کیفی بر اساس سنجش استاندارد، برخی سازمان‌های رتبه‌بندی، نهادهای دولتی و مدیران دانشگاه، تعداد نشریات علمی تولیدشده توسط اعضای هیات علمی کالج‌ها و دانشگاه‌ها را مشخص و نیز تعیین کرده‌اند که چه تعداد از این انتشارات از بانک اطلاعاتی تبلیغات برای گردآوری این اطلاعات استفاده کرده‌اند. [۳۴] در اینجا موردی وجود دارد که استانداردسازی اطلاعات می‌تواند کیفیت آن را تنزل دهد.

مشکل اول این است که این پایگاه‌های اطلاعات داده اغلب غیرقابل اعتمادند: آنها برای اندازه‌گیری خروجی در علوم طبیعی طراحی شده‌اند و اغلب اطلاعاتی تحریف‌شده در علوم انسانی و اجتماعی ارائه می‌دهند. پژوهش‌های جدید در علوم طبیعی و در برخی موارد علوم رفتاری، به صورت مقالات در مجلات کارشناسی و بازبینی‌شده منتشر می‌شود. اما در حوزه‌هایی مانند تاریخ، کتاب‌ها همچنان سرآمد انتشارات هستند و اندازه‌گیری تعداد مقالات چاپ‌شده در این حوزه تصویری تحریف‌شده ارائه می‌دهد و این تنها آغاز مشکل است.

هنگامی که اعضای هیات علمی یا تمامی بخش‌ها با تعداد نشریات‌شان -چه در قالب مقاله و چه به صورت کتاب- داوری می‌شوند، انگیزه به سمت تولید نشریات بیشتر و نه بهتر می‌رود. کتاب‌های مهم نیازمند سال‌ها پژوهش و نگارشند اما اگر سیستم تشویقی به سرعت و حجم خروجی پاداش دهد، نتیجه احتمالاً کاهش تولید آثار برجسته و در خور توجه خواهد بود. این دقیقاً همان چیزی است که علی‌الظاهر در بریتانیای کبیر در نتیجه ارزیابی پژوهش‌ها اتفاق افتاده است: جریان گسترده‌ای از نشریات که نه جالبند و نه خواننده می‌شوند. [۳۵] این مساله محدود به علوم انسانی نیست، در علوم دیگر نیز ارزیابی فقط با ارزیابی عملکرد منجر به یک جهت‌گیری به سمت انتشارات کوتاه‌مدت می‌شود تا ایجاد ظرفیت پژوهش بلندمدت. [۳۶]

در دانشگاه نیز مثل هر جای دیگر، آنچه اندازه‌گیری شود به بازی گرفته می‌شود. متناسب با «ضریب تاثیر اندازه‌گیری»<sup>۱</sup> عمل کنید. هنگامی که مشخص شد همه مقالات منتشرشده از اهمیت برابری برخوردار نیستند، تکنیک‌هایی برای سنجش تاثیر هر مقاله ایجاد شد. این تکنیک‌ها دو دسته بودند: تعداد دفعاتی که به یک مقاله در گوگل اسکولار<sup>۲</sup> یا پایگاه داده‌های تبلیغاتی ارجاع داده شده بود و

1. Impact factor measurement

2. Google Scholar

گوگل اسکولار یک موتور جست‌وجوی اینترنتی است که به طور خاص مطالب و منابع علمی را جست‌وجو می‌کند و به کاربران اجازه می‌دهد مقاله‌های مجلات، نقل‌قول‌ها، پایان‌نامه‌ها، نسخه‌های چاپی و کتاب‌های در دسترس را در وب جست‌وجو کنند. م

توجه به «ضریب تاثیر» نشریه‌ای که مقاله در آن به چاپ رسیده است؛ عاملی که به نوبه خود با توجه به فراوانی مقالات چاپ‌شده که در پایگاه‌های داده به آنها ارجاع داده‌شده، تعیین می‌شود. (البته، این روش نمی‌تواند تفاوت میان ارجاعاتی مانند نمونه‌های زیر را مشخص کند: «کتاب روشن‌گر و جامع جری زد. مولر با عنوان استبداد معیارها به طور موثری گاوهای مقدس بسیاری از سازمان‌ها را سلاخی کرد» و «درک ضعیف و سطحی جری زد. مولر او را شایسته نادیده گرفتن از سوی مدیران و دانشمندان علوم اجتماعی می‌کند.» از منظر تاثیر جدول‌بندی این دو جمله یکسان است). مجلات با توجه به رشته‌ها دسته‌بندی می‌شدند و برای بیشتر اهداف تنها ارجاعات مرتبط با رشته نویسنده در مجلات محاسبه می‌شد؛ این مساله نیز بسیار مشکل‌ساز بود زیرا علایق میان‌رشته‌ای را نادیده می‌گرفت.

علاوه بر این، در نمونه دیگری در قانون کمبل (که در فصل یک به آن پرداختیم) در تلاش برای افزایش تعداد ارجاعات خود، برخی پژوهشگران حلقه‌های ارجاعات غیررسمی تشکیل دادند و هدف اعضای آن استناد هر چه بیشتر به کار یکدیگر است. بعضی مجلات با رتبه پایین در تلاش برای بهبود «ضریب تاثیر» خود به دنبال نویسندگان مقالات پذیرفته‌شده بودند که ارجاعات بیشتری به مقالات چاپ‌شده در مجله داشته باشند. [۳۷]

سوال این است که جایگزین برای شمارش تعداد نشریات، زمان‌هایی که به آنها استناد شده و دسترسی به نشریاتی که مقالات در آنها چاپ شده چگونه است؟ پاسخ داوری حرفه‌ای است. در یک واحد دانشگاهی، ارزیابی بهره‌وری هیات علمی می‌تواند توسط استاد کرسی یا کمیته‌ای کوچک و در صورت لزوم با مشورت دیگر اعضای هیات علمی و بر اساس دانش خود، انبان تجربیات و موارد مهم در کتاب یا مقاله انجام شود. در مورد تصمیمات مهم مانند حفظ و ارتقا در رتبه‌بندی، از پژوهشگران در میان نامزدهای حوزه تخصصی خواسته می‌شود تا ارزیابی محرمانه و بسیار دقیق‌تر از کارشناسی انجام دهند. ممکن است در این فرآیند

از اعداد گردآوری شده از پایگاه‌های داده استنادی استفاده شود اما برای ارزیابی ارزش آنها اعداد نیازمند داوری مبتنی بر تجربه هستند. این داوری (قضاوت) که بر اساس تجربیات حرفه‌ای است دقیقا با اتکای بسیار زیاد به شاخص‌های عملکرد استاندارد از بین می‌رود. [۳۸] همان‌گونه که یکی از متخصصان درباره استفاده و سوءاستفاده از رتبه‌بندی گفته است: «اغلب سیستم‌های رتبه‌بندی روشی ارزان و ناکارآمد برای ارزیابی دانشمندان هستند. این عمل نه تنها منجر به ارزیابی نادرست می‌شود بلکه دانشمندان را تشویق می‌کند تا ابتدا به دنبال ارتقای رتبه‌بندی و سپس به دنبال علم باشند. روش بهتری نیز برای ارزیابی اهمیت یک مقاله یا خروجی پژوهش یک محقق وجود دارد: آن را بخوانید.» [۳۹]

### ارزش‌ها و محدودیت‌های رتبه‌بندی

رتبه‌بندی‌های عمومی از نوعی که «یو.اس.ان. دابلیو. آر» ارائه کرده، مزایای خوبی هم دارد. این رتبه‌بندی‌ها دست‌کم برای افراد ناآگاه به ارائه برخی شواهد اولیه از وضعیت نسبی نهادهای مختلف می‌پردازد. آنها کالج‌ها و دانشگاه‌ها را ترغیب می‌کنند تا اطلاعات مربوط به ابزارهای ممکن برای دانشجویان بالقوه مانند میزان نگهداری و فارغ‌التحصیلی را منتشر کنند. آنچه موجب کوتاهی در ارائه اطلاعات می‌شود توضیح و چرایی درباره بالا یا پایین بودن رتبه نگهداری و فارغ‌التحصیلی است. کالژی که دانش‌آموزان با آمادگی بالا را پذیرش می‌کند تمایل به رتبه بالای نگهداری و فارغ‌التحصیلی دارد اما برای موسسه‌هایی که هدف آنها پذیرش دانش‌آموزان با آمادگی کمتر است، معیارهای «شفاف» موجب ناکامی می‌شود زیرا با توجه به دانش‌آموزانی که پذیرش کرده‌اند می‌توانند به طور نسبی موفق شوند. دانش‌آموزان آنها به احتمال زیاد به دوره‌های تکمیلی نیاز دارند، احتمال کمتری وجود دارد که مدرکی کسب کنند و در بازار کار نیز احتمالا عملکرد ضعیف‌تری خواهند داشت. همان‌طور که بیمارستان‌های مناطق فقیرنشین به دلیل نرخ نسبتا بالای پذیرش بیماران جریمه می‌شوند (که در فصل نهم به بررسی

این موضوع می‌پردازیم)، کالج‌هایی که به افراد کم‌درآمد نیز خدمات ارائه می‌دهند احتمالاً به دلیل برخورد با جمعیتی که به آنها خدمات می‌دهند مشمول جریمه می‌شوند. رتبه‌بندی‌ها، دانشگاه‌ها را تشویق می‌کند تا تبدیل شوند به آنچه که رتبه‌بندی‌ها اندازه‌گیری می‌کنند. در واقع توجه به سمت آنچه اندازه‌گیری می‌شود، معطوف است. این امر به همگن شدن می‌انجامد زیرا مأموریت‌های متمایز و ویژه رها می‌شود و رقبا شبیه به یکدیگر خواهند شد. [۴۰]

### دسته‌بندی کالج‌ها: کارت امتیاز

وزارت آموزش در میان سنگر معیارها در ایالات متحده آمریکا قرار دارد که تحت نظارت روسای جمهور جمهوری خواه و دموکرات است. در دوره دوم ریاست جمهوری اوباما وزارت آموزش او یک «سیستم رتبه‌بندی موسسات پس از آموزش متوسطه<sup>۱</sup>» ایجاد کرد. بنا بر این بود که همه کالج‌ها و دانشگاه‌ها رتبه‌بندی شود، داده‌های آنها با توجه به «جنسیت، نژاد، قوم و دیگر متغیرها»، تفکیک و احتمالاً بودجه‌های فدرال نیز در این رتبه‌بندی‌ها در نظر گرفته و روی مواردی مانند دسترسی، مقرون به صرفه بودن و خروجی‌ها شامل انتظار درآمد پس از فارغ‌التحصیلی تمرکز شود. تد میچل<sup>۲</sup>، معاون وزیر آموزش گفت: «مردم باید بدانند که دانشجویان چگونه در موسسات دریافت‌کننده بودجه فدرال عمل می‌کنند و این عملکرد باید هنگامی که ما سرمایه‌گذاری‌ها و اولویت‌های خود را ارزیابی می‌کنیم، مورد توجه قرار گیرد.» «همچنین ما باید اقدام به ساخت مدارس کنیم که پیشرفت به سوی اهدافمان را سرعت بخشد؛ اهدافی مانند فارغ‌التحصیلی دانش‌آموزان کم‌درآمد و کاهش هزینه‌های تحصیل.» [۴۱] با برنامه‌های دولت در یک سیستم جامع رتبه‌بندی از سوی کالج‌ها و کنگره مخالفت شد و در نهایت، وزارت

1. Postsecondary Institution Ratings System

2. Ted Mitchell

آموزش یک نسخه اصلی از «کالج اسکورکارد» (کارت امتیاز کالج)<sup>۱</sup> را در سپتامبر ۲۰۱۵ منتشر کرد.

این برنامه، محصول نیات خوب و برای رفع مشکلات واقعی در ارائه آموزش عالی در نظر گرفته شده بود. یکی از این مشکلات، سابقه بسیار منفی سازمان‌های انتفاعی بود که آموزش‌های شغل‌محور در حوزه‌هایی همچون هنر آشپزی، تعمیر اتومبیل یا کمک‌های بهداشتی ارائه می‌دادند و به سرعت هم در حال گسترش بودند. بعضی از این شرکت‌ها (مانند کورینتیان<sup>۲</sup> و آی‌تی‌تی<sup>۳</sup> که هر دو در نهایت توسط دولت تعطیل شدند) با هر استانداردی غارت‌گر بودند و دانش‌آموزان بالقوه ناآگاه را شکار می‌کردند و به آنها وعده مدارکی می‌دادند که بتوانند با آن در مشاغلی پرسود مشغول به کار شوند. در حقیقت، کیفیت آموزش اغلب ضعیف بود و دانش‌آموختگان موفقیت کمی در بازار کار داشتند. علاوه بر این، حدود ۹۰ درصد شهریه از وزارت آموزش به صندوق شرکت‌های انتفاعی واریز می‌شد؛ وام‌هایی که قرار بود توسط دانشجویان وام‌گیرنده بازپرداخت شود. اما بخش انتفاعی در واکنش به یک مشکل اصلی در سطوح پایین این بخش، به مطالبات گسترده‌ای که پیامدهای آن مربوط به همه کالج‌ها و دانشگاه‌ها بود، پاسخ داد.

آنچه طرفداران معیارهای پاسخگویی بیشتر دولت از آن غافلند مشکل افزایش هزینه‌های کالج‌ها و دانشگاه‌هاست که تا حدی ناشی از گسترش کادری مدیریت است که بسیاری از آنها موظف به رعایت و عمل به موازین دولت هستند. یکی از تاثیرات قابل پیش‌بینی برنامه جدید، افزایش هزینه‌های مدیریت است؛ این امر هم با انحراف زمان بیشتر اعضای هیات علمی از تدریس و پژوهش به سوی پر کردن فرم‌های مربوط به جمع‌آوری داده و هم با افزایش تعداد مدیران برای جمع‌آوری این فرم‌ها و تامین مواد خام برای معیارهای دولت انجام می‌شود.

1. College Scorecard

یک وبسایت است که در آن اطلاعات پنج سال اخیر کالج‌های ایالات متحده آمریکا را بر اساس جنسیت، هزینه، مزایا، نرخ فارغ‌التحصیلی، میانگین وام دانشجویی و نرخ اشتغال دسته‌بندی و مقایسه می‌کند. م

2. Corinthian

3. ITT

بعضی از اهداف پیشنهادی طرح اصلی (سیستم رتبه‌بندی موسسات آموزش پس از متوسطه) نیز واقعا منحصر به فرد، در حالی که برخی دیگر کاملاً بی‌معنا و پوچ بودند. مثلاً هدف افزایش نرخ فارغ‌التحصیلان کالج با افزایش دسترسی در تضاد بود، زیرا احتمال اینکه دانشجویان با امکانات کمتر نه تنها از نظر مالی فقیرتر که آمادگی پایین‌تری هم داشته باشند، زیاد است. هرچه دانشجویان آمادگی بیشتری داشته باشند، احتمال اینکه به موقع فارغ‌التحصیل شوند نیز بیشتر است. بنابراین کالج‌ها و دیگر موسسه‌هایی که دسترسی بیشتری را برای افرادی که آمادگی کافی ندارند، ارائه می‌دهند به دلیل نرخ پایین فارغ‌التحصیلی‌شان تنبیه خواهند شد. هرچند آنها می‌توانند به دو طریق با اعداد بازی کنند؛ آنها می‌توانند -البته به قیمت دسترسی- استانداردها را برای دانشجویان ورودی افزایش دهند و به این ترتیب احتمال فارغ‌التحصیلی آنها را نیز بیشتر کنند. آنها همچنین می‌توانند با پایین آوردن استانداردها برای فارغ‌التحصیلی پاسخ دهند؛ البته با هزینه کردن کیفیت آموزش و ارزش بازاری مدرک. این احتمال نیز وجود دارد که دانشجویان بیشتر از نظر اقتصادی و شناختی پذیرش شوند و نه از نظر آمادگی علمی و به این ترتیب احتمال فارغ‌التحصیلی آنها با صرف هزینه‌های زیاد ممکن است؛ البته این مورد هم با هدف دیگر وزارت آموزش، یعنی پایین آوردن هزینه‌های آموزش مغایرت دارد.

معیار دیگری که کالج‌ها و دانشگاه‌ها ارائه می‌کردند میانگین درآمد دانشجویان پس از فارغ‌التحصیلی بود. این امر برای موسسه‌های انتفاعی و شغلی متمرکز معنا پیدا می‌کند که همان‌طور که دیدیم مستعد سوءاستفاده و دادن وعده به دانشجویان برای فارغ‌التحصیلی و مدارکی با کیفیت مشکوک است. اما برای اکثر کالج‌ها و دانشگاه‌ها نه تنها گردآوری این اطلاعات بسیار پرهزینه و در عین حال غیرقابل اعتماد است بلکه خطر تحریف شدن نیز وجود دارد. برای بسیاری از دانش‌آموزان ممتاز، آموزش حرفه‌ای به هر نحو، به‌گونه‌ای است که این اطمینان را به آنها می‌دهد که درآمد آنها حداقل تا زمانی که مشغول درس

خواندن هستند، پایین خواهد بود. بنابراین دانش‌آموخته‌ای که بلافاصله پس از فارغ‌التحصیلی در و المارت مورد استقبال قرار می‌گیرد، نمره بالاتری نسبت به فارغ‌التحصیلی دارد که به دانشکده پزشکی رفته است اما اعدادی برای نشان دادن و بنابراین «پاسخگویی» وجود دارد.

همچنان مشکل افزایش هزینه‌های تحصیل در کالج وجود دارد، هزینه‌هایی که افزایش آنها فراتر از سطح تورم است. موضوع مقرون به صرفه بودن با گرایش بسیاری از دولت‌ها به کاهش حمایت مالی از کالج‌های دولتی شدت گرفت. شاید کمترین عنصر شفافیت درباره هزینه‌های کالج، هزینه واقعی برای حضور در یک موسسه مشخص به دلیل شکاف میان قیمت برچسب و قیمت خالص باشد. قیمت برچسب هزینه رسمی شهریه، خانه و خوراک است؛ قیمت خالص مبلغ واقعی پرداخت‌شده توسط دانشجویان و والدین‌شان پس از محاسبه کمک‌های مالی بر اساس نیاز اقتصادی یا شایستگی‌های علمی است. این اختلاف اغلب قابل توجه و از نظر بسیاری خلاف واقع است: زیرا معتبرترین موسسه‌ها تمایل دارند تا بخشنده‌ترین نیز باشند؛ از آنجا که آنها کمک‌هزینه‌های بسیاری را در اختیار دارند تلاش می‌کنند تا هزینه‌های دانشجویان کارشناسی را پرداخت کنند. بنابراین، دانشجویی فقیر از نظر اقتصادی اما غنی در تعهد، ممکن است هزینه‌های واقعی رفتن به یک کالج بهتر را کمتر از کالج‌هایی بیابد که از نظر اسمی ارزان‌تر و کم‌اعتبارتر هستند. تا جایی که رتبه‌بندی‌ها چنین اطلاعاتی را منتقل می‌کنند - و کالج اسکورکارد تلاش می‌کند - آنها خدماتی واقعی ارائه می‌دهند.

با توجه به هدف اعلام‌شده اوپاما در کمک به دانشجویان و والدین‌شان برای رسیدن به «بهترین تحصیل با کمترین هزینه»، برنامه اسکورکارد روی سه معیار تاکید می‌کند: نرخ فارغ‌التحصیلی، میانگین هزینه سالانه و میزان حقوق ده ساله پس از ورود به کالج و نه بلافاصله پس از فارغ‌التحصیلی. [۴۲] این ارقام مشکل‌ساز بودند، زیرا تنها شامل داده‌های دانشجویانی بودند که از بودجه فدرال کمک هزینه دریافت

می‌کردند؛ یعنی این نتایج در مورد دانشجویانی کاربرد دارد که دارای پیشینه اقتصادی ضعیف‌تری هستند. احتمالاً والدین ثروتمندتر دارای درآمد بیشتر [۴۳] و ارقام حقوق مبهم هستند، هرچند این ارقام در کالج‌های مختلف تحت مدیریت‌های متفاوت و بسته به پیشینه‌های مختلف مجموعه دانش‌آموزان متفاوت است. از همه نگران‌کننده‌تر این حقیقت است که کالج اسکورکارد هیچ تلاشی نمی‌کند در جداسازی سهم مدرسه در درآمد از آنچه که به طور منطقی می‌توان انتظار داشت بر اساس درآمد خانواده، نمرات آزمون دانش‌آموزان یا سطح مدارکی که ارائه می‌دهند باشد. [۴۴] با این حال، خروجی‌های کالج به شدت با ورودی‌ها مرتبط است: دانشجویانی که با سطوح بالاتر علمی وارد کالج می‌شوند - و بیشتر اوقات دارای والدینی با سطوح بالای آموزشی و درآمدی هستند - در ارزیابی‌های استاندارد خروجی‌های کالج نیز موفق‌ترند. [۴۵]

نهاد بروکینگز<sup>۱</sup> تلاش کرده است تا با استفاده از اطلاعات اضافی بر این مانع غلبه و «ارزش افزوده» را محاسبه کند که به معنای افزایش درآمد ارائه‌شده توسط هر کالج با توجه به داده‌های موجود و پیشینه دانشجویان ورودی هر موسسه است. امید می‌رود که چنین معیارهایی «برای بسیاری از افراد علاقه‌مند به دانستن اینکه کالج‌های خوب (خاص) چگونه دانشجویان را برای مشاغل پردرآمد آماده می‌کنند، مفید باشد.»

### پیام معیارها: کالج‌ها برای کسب درآمد هستند

بیا بید صحت و اعتبار این معیارها را برای کشف یک مساله مهم‌تر کنار بگذاریم: پیامی که توسط خود معیارها منتقل می‌شود. کالج اسکورکارد تحصیلات کالج را تنها از نظر اقتصادی در نظر می‌گیرد: تنها دغدغه آن بازگشت سرمایه و فهم آن به عنوان ارتباط میان هزینه‌های پولی

1. The Brookings Institution

یک سازمان غیرانتفاعی بی‌طرف آمریکایی است که در زمینه تحلیل مسائل سیاسی، اجتماعی و اقتصادی فعالیت می‌کند. این نهاد در سال ۱۹۱۶ میلادی تاسیس شده، یکی از قدیمی‌ترین موسسه‌ها در نوع خود و مقر آن در شهر واشنگتن است. (منبع: ویکی‌پدیای فارسی)

کالچ و افزایش درآمد است که در نهایت منتهی به کسب یک مدرک می‌شود. البته اینها ملاحظات قانونی هستند: هزینه‌های کالچ درصد زیادی از درآمد خانواده را از بین می‌برد یا دانشجو را درگیر وام و بدهی می‌کند، و در نهایت درآمدزایی را به منزله یکی از مهم‌ترین وظایف زندگی مطرح می‌کند. اما این تنها وظیفه در زندگی نیست؛ این تنها یک درک ضعیف از تحصیل در کالچ است که آن را صرفاً از نظر توانایی در کسب و افزایش درآمد در نظر می‌گیرد. [۴۶] با این حال، ایده‌آل آموزش این است که کالچ اسکورکارد را مانند معیارهای مشابه تشویق و ترغیب کند. اگر پرورش را (که به تولید و بقا گرایش دارد) از آموزش (که به بقای معنادار متمایل است) جدا کنیم، کالچ اسکورکارد فقط در مورد گزینه اول است. [۴۷] در واقع سیستم‌های اسکورکارد و بروکینگز به دنبال رتبه‌بندی بیشتر موسسه‌هایی هستند که بر حوزه‌های مهندسی و فناوری و مواد تولیدی تمرکز دارند. یک نوع رضایت از زندگی که ناشی از یک دوره تاریخ هنر است و به شما اجازه می‌دهد که یک کار هنری را در زمینه تاریخی آن درک کنید؛ یا یک دوره موسیقی که شما را برای شنیدن مضمون و متغیرهای یک سمفونی یا تعبیر جاز با یک صدای استاندارد پرورش می‌دهد؛ یا یک دوره ادبیات که فهم و علاقه شما به شعر را افزایش می‌دهد؛ یا یک دوره اقتصاد که به شما در فهم نهادهای اقتصادی مهم کمک می‌کند؛ یا یک دوره زیست‌شناسی که چشمان شما را بر روی شگفتی‌های ساختار بدن انسان باز می‌کند؛ هیچ یک از این دوره‌ها با معیار بازگشت سرمایه تسخیر نشده‌اند. این واقعیت نیز درست نیست که کالچ اغلب مکانی برای شکل‌گیری دوستی‌های مادام‌العمر و مهم‌ترین آنها یعنی ازدواج است. همه این موارد باید در «بازگشت سرمایه» مورد بررسی قرار گیرند: اما از آنجا که از نظر کمی قابل اندازه‌گیری نیستند، به آنها توجهی نمی‌شود.

خطر معیارهایی که صرفاً بر بازگشت پولی سرمایه تمرکز دارند این است که مانند بسیاری از معیارها، آنها نیز بر رفتار اثرگذارند. در حال حاضر، دانشگاه‌ها که در رأس این رتبه‌بندی‌ها قرار دارند بخش بزرگی

از فارغ‌التحصیلان خود را به سوی مشاغل پردرآمدی مانند بانکداری، مشاوره و شرکت‌های حقوقی سطح بالا سوق می‌دهند. [۴۸] همه اینها مشاغل شریفی هستند اما آیا این واقعا به نفع ملت است که بهترین و درخشان‌ترین آنها تشویق به انتخاب چنین مشاغلی شوند؟ یکی از تاثیرات قابل پیش‌بینی این کار مرتبط است با میزان درآمد آینده فارغ‌التحصیلان با رتبه‌بندی کالج‌ها، که همین امر خود موجب ایجاد انگیزه برای موسسات است تا دانشجویان خود را به ورود در این مشاغل پردرآمد تشویق و ترغیب کنند. کالج‌هایی که فارغ‌التحصیلان‌شان وارد مشاغلی کم‌درآمد مانند تدریس یا خدمات عمومی می‌شوند، تنبیه خواهند شد. [۴۹]

جامعه سرمایه‌داری برای شکوفایی خود وابسته به نهادهایی است که وزنه‌ای در برابر بازار با تمرکز بر سود پولی ایجاد کند. آماده‌سازی دانش‌آموزان و دانشجویان دانشگاه برای ایفای نقش‌های خود به عنوان شهروندان، دوستان، همسران و مهم‌تر از همه آماده‌سازی آنها برای یک زندگی دارای غنای فکری جزو نقش‌ها و وظایف کالج‌ها و دانشگاه‌هاست. انتقال مهارت‌های قابل فروش نیز نقش دیگر این نهادهاست اما مطیع ساختن تمام و کمال آموزش عالی برای ایجاد ظرفیت کسب درآمد آتی عبارت است از مقید ساختن آن به اندازه‌گیری با یک معیار به شدت نادرست.

## فصل ۸

# مدارس

تلاش برای دستیابی به نتایج قابل اندازه‌گیری در سیاست‌گذاری دولت برای کلیه مقاطع تحصیلی مدرسه، حتی بسیار مهم‌تر بوده است. دیان راویچ<sup>۱</sup>، مورخ آموزش (و مقام پیشین وزارت آموزش) می‌گوید: «فرمانداران، مدیران ارشد اجرایی شرکت‌ها، دولت اول بوش و دولت کلینتون در یک مورد توافق داشتند: آنها به دنبال نتایج قابل اندازه‌گیری بودند، آنها می‌خواستند دلارهای مالیاتی سرمایه‌گذاری شده در آموزش عمومی، بازده خوبی داشته باشد.» [۱] در بخش عمومی نمایش اسب معیارها تبدیل به «هیچ کودکی نباید از تحصیل عقب بماند» شد؛ قانونی که در زمان جرج دبیلوبوش در سال ۲۰۰۱ با حمایت دو حزب به تصویب رسید و عنوان رسمی آن عبارت بود از: «قانونی برای کاهش شکاف میان موفقیت و پاسخگویی، انعطاف‌پذیری و انتخاب، چنانکه هیچ کودکی از تحصیل عقب نماند.»

### مشکل و راه حل پیشنهادی

این برنامه به دنبال حل یک مشکل واقعی بود: به‌رغم تلاش‌های

1. Diane Ravitch

اساسی در سطح دولت برای یکسان‌سازی هزینه‌ها در میان مدارس مناطق، یک اختلاف دائمی در عملکرد مدارس گروه‌های قومی وجود داشت. طرفداران این اصلاحات اظهار کردند که این قانون با پاسخگو نبودن آموزگاران و مدیران مقابله می‌کند و برای بهبود نتایج با نظم بخشیدن به رفتار آموزگاران، دانش‌آموزان و مدارس مطابق با «اهداف عملکرد سیستم»، انگیزه ایجاد می‌کند. [۲] تصور می‌شد که مقصر این وضعیت، حرفه‌ای نبودن آموزگاران مدارس دولتی است.

این قانون نتیجه یک دهه لابی‌گری سنگین یک ائتلاف فوق‌العاده ناهمگن است: گروه‌های تجاری نگران کیفیت نیروی کارند؛ گروه‌های حقوق مدنی که نگران دستاورد گروه‌های مخالفند و اصلاح‌طلبان آموزشی ناخشنود از ناکامی آموزش مدارس دولتی خواستار آزمون‌ها و ارزیابی‌های استاندارد ملی هستند. [۳] مزایای چنین اقداماتی (اندازه‌ها) از منظر آرمان‌خواهی بسیار اندک بود. ویلیام کولبرگ<sup>۱</sup> از اتحاد ملی کسب‌وکار<sup>۲</sup> گفت که ایجاد سیستم استاندارد ملی همراه با ارزیابی اطمینان می‌دهد که هر دانش‌آموز فقط با بروز تواناییش در خواندن، نوشتن، حساب و شرکت در کلاس‌هایی در سطوح جهانی در موضوعات مدارس عمومی می‌تواند مدرسه را ترک کند. [۴]

اولین ثمره این تلاش در سطح فدرال بهبود قانون مدارس آمریکا بود که در زمان ریاست جمهوری بیل کلینتون در سال ۱۹۹۴ تصویب شد. در همین حال، جرج دابلیو بوش فرماندار تگزاس بدل به قهرمان آزمون‌های اجباری و پاسخگویی آموزشی شد. طبق این قانون که در اوایل دوران ریاست جمهوری بوش تصویب شد: هیچ کودکی نباید از تحصیل عقب بماند، ایالت‌ها باید سالانه از هر دانش‌آموز در کلاس‌های سه تا ۸ در درس ریاضی، خواندن و علوم امتحان بگیرند. این قانون برای این بود که همه دانش‌آموزان را تا سال ۲۰۱۴ به «مهارت علمی» برساند و اطمینان حاصل کند که همه گروه‌های دانش‌آموزی - از جمله

1. William Kolberg

2. National Alliance of Business

سیاه‌پوستان و اسپانیایی‌زبانان که برای ارزیابی مقایسه‌ای جدا شده بودند- در هر مدرسه «پیشرفت مناسب سالانه‌ای» دارند و هر سال مهارت بیشتری کسب می‌کنند. این قانون مجموعه‌ای از مجازات‌ها و تنبیهات را برای مدارس وضع کرد که در آنها برخی گروه‌های دانش‌آموزی پیشرفت کافی نداشته‌اند. این قانون تحت حمایت سنا نیز بود. ادوارد کندی به‌رغم مخالفت برخی دموکرات‌های لیبرال و جمهوری خواهان محافظه‌کار مخالف گسترش قدرت فدرال، این قانون را با حمایت جمهوری خواهان و دموکرات‌ها به تصویب رساند.[۵]

با وجود گذشت بیش از یک دهه از اجرای این قانون، مزایای پاسخگویی آن همچنان مبهم است (دیگر جوانب این قانون، افزایش انتخاب مدارس، ایجاد مدارس خصولتی<sup>۱</sup> و صلاحیت‌های بیشتر آموزگاران بوده است؛ به نظر می‌رسد این قانون در این حوزه‌ها موفق بوده اما این موارد فراتر از موضوع مورد بحث ماست). طرفداران این قانون هرگونه مدرکی دال بر پیشرفت در هر آزمون، در هر کلاس و در هر گروه جمعیتی را نشانه اثربخشی آن می‌دانند. اما نمرات آزمون برای دانش‌آموزان دبستان صرفاً به مقدار ناچیزی افزایش یافت و نه بیشتر از مقداری که پیش از وضع قانون بود و تاثیر آن بر نمرات دانش‌آموزان دبیرستانی بسیار محدودتر بود.

اثر اصلی این قانون توجه بیشتر به «شکاف موفقیت» بود؛ تفاوت در عملکرد تحصیلی دانش‌آموزان آسیایی، اسپانیایی‌زبانان، سفیدپوستان و سیاه‌پوستان.[۶] آسیایی‌ها از سفیدپوستان نمرات بیشتری می‌گرفتند و سفیدپوستان نیز به نوبه خود عملکرد بهتری از سیاه‌پوستان و اسپانیایی‌زبانان داشتند. از همه مهمتر ضعف همیشگی دانش‌آموزان آمریکایی آفریقایی تبار بود. هشت سال پس از معرفی این قانون، نمرات دانش‌آموزان تغییری نکرد و میانگین نمرات در امتحانات ملی آزمون‌های پیشرفت تحصیلی ارزیابی ملی در دروس انگلیسی و ریاضیات

1. Charter School

مدارسی در آمریکا که با وجود دریافت کمک هزینه از بودجه دولتی اما به طور مستقل اداره می‌شوند. هنوز بر سر اینکه این مدارس خصوصی یا دولتی هستند توافق نظر وجود ندارد. م

برای دانش‌آموزان هفده ساله در حد فاصل سال‌های ۱۹۷۰ تا ۲۰۰۸ تقریباً بدون تغییر باقی ماند. در واقع، نمرات برای هر گروه (آسیایی، سفیدپوست، سیاه‌پوست و اسپانیایی‌زبان) تا اندازه‌ای بالا رفت اما به دلیل تغییر ترکیبات قومی دانش‌آموزان (به‌ویژه افزایش درصد دانش‌آموزان اسپانیایی‌زبان که نمرات کمتری نسبت به هم‌تایان آسیایی یا سفیدپوست خود می‌گرفتند) میانگین نمرات ملی ثابت ماند. [۷]

### پیامدهای ناخواسته

پیامدهای ناخواسته برنامه آزمون و پاسخگویی هیچ کودکی نباید عقب بماند، ملموس است و بسیاری از نقایص ثبات معیار را نشان می‌دهد. بر اساس این برنامه نمرات در آزمون‌های استاندارد، معیاری عددی هستند که موفقیت و شکست بر مبنای آنها قضاوت می‌شود. خطر این امر برای آموزگاران و مدیرانی که مشاغل آنها و افزایش حقوق‌شان بستگی به این شاخص عملکرد دارد، بالاست. بنابراین تعجیبی ندارد که آموزگاران (با تشویق مدیران‌شان) زمان کلاس‌های خود را به تدریس درس‌های آزمودنی - ریاضی و انگلیسی - اختصاص دهند و از دیگر موضوعات مانند تاریخ، مطالعات اجتماعی، هنر، موسیقی و تربیت بدنی غفلت کنند. آموزش ریاضی و انگلیسی به جای فرآیندهای شناختی گسترده‌تر، بر مهارت‌های مورد نیاز آزمون تمرکز دارد: دانش‌آموزان اغلب استراتژی‌های آزمون دادن را یاد می‌گیرند تا دانش بنیادین را. همان‌طور که در سریال «شنود» شبکه اچ‌بی‌او نشان داده شد، زمان زیادی از کلاس صرف تمرین برای آزمون‌ها می‌شود که به‌سختی منبع تحریک و تشویق دانش‌آموزان است. همان‌گونه که متیو آرنولد<sup>۱</sup> یک قرن و نیم قبل پیش‌بینی کرده بود: از آنجا که دانش‌آموزان زبان انگلیسی را به‌گونه‌ای یاد می‌گیرند که بتوانند با خواندن متون کوتاه به پرسش‌های چندگزینه‌ای پاسخ دهند، آنها در خواندن متون بلند و نوشتن مقالات بسیار ضعیف خواهند بود. [۸]

1. Mathew Arnold

مشکل مربوط به استفاده از آزمون‌های استاندارد که به‌درستی طراحی شده‌اند و می‌توانند معیارهای سودمندی برای سنجش توانایی و پیشرفت دانش‌آموزان باشند، نیست. آزمون ارزش افزوده که تغییرات عملکرد دانش‌آموزان را هر ساله اندازه‌گیری می‌کند، واقعاً سودمند است. این آزمون‌ها کمک می‌کنند که عملکرد ضعیف آموزگاران مشخص شود و آنها از سیستم حذف شوند. [۹] مهم‌تر از همه، آزمون ارزش افزوده می‌تواند در حکم یک ابزار تشخیصی مفید واقع شود، زیرا توسط آموزگاران مورد استفاده قرار می‌گیرد تا دریابند کدام جنبه‌های برنامه درسی موثرترند. اما آزمون‌های ارزش افزوده زمانی بیشترین تاثیر را خواهند داشت که «میزان ریسک پایین باشد». [۱۰] تاکید بر این آزمون‌ها به منزله معیار مهمی در ارزیابی مدارس، موجب انگیزه‌های نامعقول می‌شود: تمرکز بر آزمون‌ها به قیمت از دست رفتن اهداف بزرگ‌تر موسسه است.

آزمون‌های زیاد موجب اختلال‌های بیشتر نیز می‌شود، از جمله کار سطحی (سرشیر گرفتن): مطالعات مدارس در تگزاس و فلوریدا نشان می‌دهد که میانگین سطوح موفقیت دانش‌آموزان با دسته‌بندی دوباره دانش‌آموزان ضعیف به عنوان افراد ناتوان افزایش می‌یابد، از این رو آنها از مجموعه ارزیابی حذف می‌شوند. [۱۱] یا کنار گذاشتن و تقلب کردن؛ آموزگارانی که پاسخ‌های دانش‌آموزان را تغییر می‌دهند یا آزمون‌هایی که احتمال دارد نمرات دانش‌آموزان در آنها پایین باشد را کنار می‌گذارند؛ پدیده‌ای که در آتلانتا، شیکاگو، کلیولند، دالاس، هیوستون، واشنگتن دی.سی و دیگر شهرها ثبت شده است. [۱۲] یا شهرداران و فرمانداران که با کاهش سطح دشواری آزمون‌ها و پایین آوردن نمرات قبولی به سوی اهداف تعیین‌شده در راستای افزایش نرخ قبولی و نشان دادن موفقیت اصلاحات آموزشی خود حرکت می‌کنند. [۱۳]

تاکید بر عملکرد اندازه‌گیری‌شده از طریق آزمون‌های استاندارد موجب ایجاد یک خروجی گمراه‌کننده دیگر نیز می‌شود؛ همان‌گونه که قانون کمبل (در فصل ۱ توضیح دادیم) پیش‌بینی کرده است: اعتبار پیش‌بینی‌کننده

آزمون‌ها از بین می‌رود. آزمون‌های عملکرد برای ارزیابی توانایی و دانشی که دانش‌آموزان در دوران تحصیل خود کسب کرده‌اند، طراحی شده است. هنگامی که آموزش به جای افزایش عملکرد دانش‌آموز بر «عملکرد دانش‌آموز در امتحانات» تمرکز می‌کند، دیگر آنچه را که برای ارزیابی آن ایجاد شده است، اندازه‌گیری نمی‌کند. مثلاً اگر زمان کلاس به تمرین پرسش‌های چهارگزینه‌ای مشابه آزمون اختصاص یابد (مثلاً استفاده از پرسش‌های آزمون‌های گذشته)، دانش‌آموزان بدون اینکه چیز زیادی درباره موضوع آزمون آموخته باشند، چه‌بسا نمرات بالاتری کسب کنند. [۱۴]

تنها چند سال پیش از تصویب قانون هیچ کودک‌کی نباید عقب بماند، دولت انگلیس نظام ارزیابی‌های معیاری خود را برای سیستم مدارس راه‌اندازی کرده بود. در سال ۲۰۰۸، یک کمیسیون پارلمانی که درباره این سیستم تحقیق می‌کرد بسیاری از اختلالات عملکردی مشاهده‌شده در ایالات متحده آمریکا را پیدا کرد. [۱۵]

### پافشاری بر داده‌ها

با وجود مشکلات روش آزمون و پاسخگویی قانون «هیچ کودک‌کی نباید عقب بماند»، وزارت آموزش دولت اوباما باز هم بر پاسخگویی و معیارهای آموزش مقاطع تحصیلی در مدارس پافشاری و در سال ۲۰۰۹ قانون «مسابقه تا اوج»<sup>۱</sup> را معرفی کرد؛ برنامه‌ای که از بودجه‌های قانون بهبود اوضاع اقتصادی و سرمایه‌گذاری مجدد<sup>۲</sup> در راستای وادار کردن ایالت‌ها به «اتخاذ ارزیابی‌ها و استانداردهای آمادگی برای کالج و شغل؛ ایجاد سیستم‌های داده که میزان رشد و موفقیت دانش‌آموزان را اندازه‌گیری کند و ایجاد ارتباط میان موفقیت دانش‌آموزان با آموزگاران و مدیران استفاده می‌کرد.» [۱۶] در حالی که قانون هیچ کودک‌کی نباید عقب بماند بر اندازه‌گیری عملکرد کلی مدارس تمرکز داشت، «مسابقه تا اوج» معیارهای عملکرد را برای آموزگاران در سطح فردی گسترش

1. Race to the Top

2. American Recovery and Reinvestment Act of 2009 (ARRA)

داد. این قانون به ایالات و مدارس منطقه که به دنبال اجرای برنامه معیارها بودند بودجه می‌داد. اکنون آموزگاران بر اساس تغییرات قابل اندازه‌گیری در پیشرفت دانش‌آموزان خود پاداش می‌گیرند. این امر در حکم «امتیاز ارزش افزوده» یا «پیشرفت دانش‌آموز» شناخته می‌شود. این قابل درک بود که معلمان نتوانند مسئولیت بالا و پایین شدن نمرات شاگردان خود را بپذیرند زیرا مشخص بود که این امر بستگی به بسیاری از عوامل بیرونی دارد که معلمان کنترلی روی آنها ندارند اما آنها باید مسئولیت اینکه دانش‌آموزان‌شان چه اندازه در طول سال آموخته‌اند را بپذیرند. تصور بر این بود که برای یافتن «ارزش افزوده» از دانش‌آموزان در آغاز و پایان سال تحصیلی آزمون گرفته و بر اساس نتیجه آنها به معلمان پاداش داده شود (هرچند این برنامه طبق عوامل ریسکی مانند نژاد و پیشینه خانوادگی تنظیم شده بود)، در برخی ایالت‌ها، نمرات ارزش افزوده، نیمی از نمره ارزیابی آموزگاران را به خود اختصاص می‌دهد. تولید داده‌های مورد نیاز برای ارزیابی معلمان در قانون «مسابقه تا اوج» نیازمند گسترش بیشتر آزمون و ارزیابی بود. [۱۷]

پذیرش معیارهای عملکرد ارزش افزوده برای معلمان با یافته‌های اقتصاددانان تشدید شد. معیارهای اولیه نشان داد که برخی آموزگاران حقیقتاً بهتر از دیگرانند و دانش‌آموزان آنها نیز موفقیت تحصیلی بیشتری کسب کرده‌اند. با استفاده از معیارهای محدود، برخی اقتصاددانان به این نتیجه رسیده‌اند که اگر ۱۵ درصد از بهترین آموزگاران به دانش‌آموزان تدریس کنند و یا آموزگارانی که ۲۵ درصد از پایین‌ترین نمرات را کسب کرده‌اند در همان سال اول اخراج شوند، شکاف میان نمرات دانش‌آموزان کمتر خواهد شد. هرچند با گذشت زمان مشخص شد که افزایش سالانه نمرات ارزش افزوده به مرور زمان کم می‌شود. [۱۸]

### پرداخت مبتنی بر عملکرد

با همان انگیزه و منطقی که به «مسابقه تا اوج» انجامید، مدارس مناطق، برنامه‌های آزمایشی پرداخت مبتنی بر عملکرد را آغاز کردند و به معلمان

مطابق معیارهای ارزش افزوده‌شان پاداش دادند. نتایج ناامیدکننده بود؛ یک آزمایش در سطح بسیار بزرگ برای پرداخت آموزگاران بر مبنای عملکردشان در شهر نیویورک سیتی بین سال‌های ۲۰۰۷ تا ۲۰۰۹ انجام شد. این مطالعه آزمایشی که توسط اقتصاددانی به نام رونالد فریر<sup>۱</sup> انجام شد، او را به این نتیجه رساند که «هیچ مدرکی مبنی بر اینکه انگیزه‌های معلمان موجب بهبود عملکرد دانش‌آموزان، فارغ‌التحصیلی و ... شود یا انگیزه‌ها رفتار معلمان و دانش‌آموزان را تغییر دهد، وجود ندارد.» [۱۹] مطالعه‌ای با نتیجه مشابه در سال ۲۰۱۱ در مرکز ملی انگیزه‌های عملکرد در دانشگاه وندربیلت انجام شد.<sup>۲</sup> این مطالعه نشان داد که دادن پاداش به آموزگاران نشویل<sup>۳</sup> بر اساس رتبه‌های ارزش افزوده‌شان تاثیر قابل توجهی نداشته است. [۲۰] مطالعات پیشین که به اواسط دهه ۱۹۸۰ بازمی‌گردند نیز به نتایج مشابهی دست یافته بودند. با وجود چنین شواهدی، باز هم باور به پرداخت مبتنی بر عملکرد به حدی قوی است که با وجود همه کاستی‌ها دائما به آن رجوع می‌شود. [۲۱]

شکست برنامه‌های پرداخت مبتنی بر عملکرد اندازه‌گیری شده در دستیابی به نتایج، دولت فدرال را از اختصاص منابع هر چه بیشتر برای چنین تلاش‌هایی متوقف ساخت. مثلا در سال ۲۰۱۰، وزارت آموزش ۶۲ برنامه را در ۲۷ ایالت انتخاب کرد تا به آنها از صندوق تشویق معلمان<sup>۴</sup> برای دوره‌های پنج ساله در حدود ۱/۲ میلیارد دلار بدهد. چنین تلاش‌هایی منحصر به ایالات متحده نیست. برنامه‌های مشابهی برای افزایش ارتباط آموزگاران، مقام آنها و ارتقای عملکرد اندازه‌گیری شده‌شان در انگلستان، پرتغال، استرالیا، شیلی، مکزیک، رژیم صهیونیستی و هند انجام شده است. [۲۲]

1. Roland Fryer

2. National Center on Performance Incentives at Vanderbilt University

3. Nashville

4. Teacher Incentive Fund

## «شکاف موفقیت» از بین نرفتنی

شاید نگرانی عمده طرفداران معیارها در حوزه آموزش آمریکا، تفاوت در دستیابی به آموزش در میان گروه‌های معین قومی و نژادی باشد. این موضوع انگیزه مهمی برای پیشگامان قانون «هیچ کودکی نباید عقب بماند» بود و در سیاست وزارت آموزش دولت اوباما همچنان مهم باقی ماند و مورد اصلاح و تجدیدنظر قرار گرفت و تحت عنوان «هر دانش‌آموزی موفق می‌شود» در اواخر سال ۲۰۱۵ تصویب شد (این عنوان همچون عناوین دیگر «هیچ کودکی نباید عقب بماند» یا «عملیات آزادی عراق»<sup>۱</sup> بسیار امیدوارکننده بود). این نگرانی محدود به سطح فدرال نبود: این موضوع در سیاست آموزشی بسیاری از ایالت‌ها و تعداد بی‌شماری از شهرداری‌ها بسیار چشمگیر بود و در دستور کار آموزگاران کالج‌ها قرار داشت. مدارس به طور فزاینده‌ای به منزله «کارخانه‌های پایان‌دهنده شکاف» شناخته می‌شدند. [۲۳]

با این وجود، جالب است پس از دهه‌ها جمع‌آوری و تبلیغ این معیارها، نتیجه کم‌وبیش بدون تغییر مانده است. وضعیت سیاه‌پوستان و اسپانیایی‌زبانان نسبت به سفیدپوستان به طور چشمگیری ثابت است. هرچند هنگامی که دانش‌آموزان در کلاس‌های ۴ و ۸ سنجش شدند، برخی نوسانات جزئی مشاهده شد اما این تغییرات در نتیجه نهایی -در پایان دوره دبیرستان یعنی کلاس ۱۲- تاثیری نداشت.

از دانش‌آموزان سراسر ایالات متحده در کلاس‌های ۴، ۸ و ۱۲ در درس خواندن و ریاضی آزمون گرفته می‌شود. این آزمون‌ها «ان.آ.ای.پی» (ارزیابی ملی پیشرفت تحصیلی)<sup>۲</sup> هستند. کارشناسان آنها را در حکم شاخص‌های نسبتاً مطمئنی از عملکرد می‌دانند زیرا برخلاف بعضی آزمون‌های دیگر آنها دارای «ریسک‌های پایین» نیز هستند: این یعنی میزان ثروت دانش‌آموزان، معلمان یا مدارس تاثیر بر نتایج ندارند و بنابراین برای معلمان انگیزه کمی وجود دارد تا امتحانات سطحی، برگزار یا با هدف امتحان، تدریس یا نتایج را دستکاری کنند. مرکز

1. Operation Iraqi Freedom

2. NAEP (National Assessment of Educational Progress)

ملی آمار آموزشی یک گزارش سالانه منتشر و در آن وضعیت و روند پیشرفت تحصیلی گروه‌های نژادی و قومی، مقایسه میزان نسبی موفقیت میان آسیایی‌ها، سفیدپوستان، اسپانیایی‌زبانان و سیاه‌پوستان (و برخی زیرمجموعه‌های این گروه‌ها) را در طول زمان ارائه کرده است. این یافته‌ها خود گویا هستند. برای کسانی که آزمون کلاس ۱۲ را گذرانده‌اند، شکاف موفقیت در خواندن میان سفیدپوستان و اسپانیایی‌زبانان در سال ۲۰۱۳ تفاوتی با سال ۱۹۹۲ نداشت (۲۲ نمره در مقیاس ۵۰۰، میانگین نمره در سال ۲۰۱۳ برابر با ۲۸۸ بود). البته، شکاف میان سفیدپوستان و سیاه‌پوستان در سال ۲۰۱۳ (۳۰ امتیاز) بیشتر از ۱۹۹۲ (۲۴ امتیاز) است. در مورد درس ریاضی، این گزارش، عملکرد نسبی هر گروه را در سال‌های ۲۰۰۵، ۲۰۰۹ و ۲۰۱۳ مقایسه می‌کند. نتیجه: شکاف نمرات میان سفیدپوستان و همسالان سیاه‌پوست و اسپانیایی‌زبان‌شان بدون تغییر باقی مانده است. [۲۴]

ناتوانی مدارس و تاثیرگذاری نسبتاً کم بر موفقیت آموزشی جای تعجبی ندارد. بنابر گزارش کلمن<sup>۱</sup> «فرصت آموزشی برابر» (۱۹۶۶) که در دولت جانسون اجرا شد مشخص کرد که خروجی مدارس وابسته به ورودی آنهاست: از این رو، عملکرد دانش‌آموزان ارتباط نزدیکی با پیشرفت اجتماعی، اقتصادی و تحصیلی والدین‌شان دارد. [۲۵] «مدارس خوب» تمایل دارند تا دانش‌آموزانی زرنگ‌تر، کنجکاو‌تر و خویشتن‌دار داشته باشند؛ اینان معمولاً فرزندان والدینی با چنین ویژگی‌هایی هستند. از آنجا که این ویژگی‌ها به موفقیت منتهی می‌شوند، این تمایل در خانواده‌ها و والدین موفق نیز وجود دارد که فرزندان‌شان را برای کسب علم به مدرسه بفرستند.

بنابراین پیشرفت‌های کلی در مدرسه منجر به برابری بیشتر نتایج نمی‌شود. همان‌طور که دانشمند علوم سیاسی ادوارد بانفیلد<sup>۲</sup> درباره نسل قبل یادآور شده است: «تمام آموزش‌ها متوجه طبقات متوسط و بالا هستند زیرا این طبقات ویژگی‌هایی تربیت‌پذیر دارند.» ارتقای کیفیت

1. Coleman Report

2. Edward Banfield

مدارس ممکن است خروجی‌های کلی آموزش را بالا ببرد اما همین امر به جای کاهش شکاف موفقیت میان کودکان از خانواده‌هایی با سطوح مختلف سرمایه انسانی آن را افزایش می‌دهد. [۲۶]

چنین خروجی‌هایی شاید منجر به این نتیجه‌گیری شود که شکاف موفقیت با تحصیل و آموزش تمام نمی‌شود و دلایل این امر فراتر از درهای مدرسه است. بنابراین، اندازه‌گیری بی‌کم‌وکاست ادامه می‌یابد. شاید این به همان دلیلی است که بانفیلد اشاره کرده است؛ این تصور که بعضی مشکلات غیر قابل حل از نظر اخلاقی برای بخش بزرگی از آمریکایی‌های تحصیل کرده قابل قبول نیست. [۲۷] وقتی شکاف‌هایی در موفقیت تحصیلی به وجود می‌آید، به نظر می‌رسد در صورت نبود پیشرفت قابل توجه در نتایج، منابع اختصاص داده‌شده به اندازه‌گیری مداوم به نشانه‌ای جدی از بی‌اخلاقی بدل می‌شود.

### هزینه‌های تلاش برای پایان دادن به شکاف

البته، نمرات آزمون‌های پیشرفت تحصیلی انگلیسی و ریاضی نمی‌تواند تمام مزایای تحصیلات دوران مدرسه را اندازه‌گیری کند. دلیل این موضوع تحریف شدن یا بی‌اهمیت بودن نمرات ارزیابی ملی پیشرفت تحصیلی نیست. این آزمون‌ها یک سنجش سودمند از دانش دانش‌آموزان در موضوعات امتحانی ارائه می‌دهد. با این حال در مدرسه چیزهای بسیار بیشتری از یادگیری انگلیسی و ریاضی وجود دارد: نه تنها دیگر موضوعات علمی بلکه تشویق علایق در جهان و تربیت و ترویج عادات رفتاری (خویش‌داری، پشتکار، توانایی همکاری با دیگران) که احتمال موفقیت در دنیای بزرگسالان را افزایش می‌دهد. ترویج و گسترش این ویژگی‌های غیرشناختی شاید در کلاس‌ها و مدارس بدون انعکاس یافتن در معیارهای عملکرد مبتنی بر نمرات آزمون تداوم داشته باشد. [۲۸]

در حقیقت، تاکید روزافزون بر آزمون گرفتن از دانش‌آموزان در درس انگلیسی و ریاضی در همان دوران مهدکودک چه‌بسا به قیمت فعالیت‌های غیرعلمی تمام شود؛ فعالیت‌هایی همچون هنرها و بازی‌های خلاقانه که

در رشد فردی نقش دارند، اما به‌سادگی اندازه‌گیری نمی‌شود. [۲۹] علاوه بر این، شاید زمانی که دانش‌آموزان، آموزگاران بهتری داشته باشند دستاوردهای علمی بیشتری نیز کسب کنند اما این دستاوردها به مرور زمان کمرنگ می‌شود و از بین می‌رود، در حالی که به نظر می‌رسد دستاوردهای غیرشناختی پابرجا باقی می‌ماند. [۳۰] مسائل مربوط به رشد شخصیت که موجب شده برخی قانون‌گذاران تلاش کنند تا سنجش شخصیت را وارد سیستم‌های پاسخگویی کنند نیز مهم است! [۳۱]

از این رو، هزینه‌های تلاش در استفاده از معیارها در راستای تبدیل مدارس به کارخانه‌های پایان‌دهنده به شکاف، صرفاً پولی نیستند. وظیفه بزرگ‌تر مدارس، تدریس تاریخ و علوم مدنی است که به دلیل توجه و تمرکز بر بهبود نمرات خواندن و ریاضی گروه‌هایی با عملکرد ضعیف مورد غفلت قرار گرفته‌اند. استراتژی‌های آموزشی که ممکن است برای دانش‌آموزان ضعیف موثر واقع شود (مانند روزهای طولانی‌تر در مدرسه و تعطیلات کوتاه‌تر در تابستان) چه‌بسا برای برخی دیگر تأثیری متضاد داشته باشد. و منابع از افراد با استعدادتر و خوش‌ذوق‌تر برای حداکثر آموختن دور شده است که ممکن است همین امر کلید عملکرد اقتصاد ملی را داشته باشد. [۳۲]

تاکید بر شکاف موفقیت و فشار برای پایان دادن به آن، اثرات نگران‌کننده دیگری نیز دارد. یکی از این آثار، سرزنش معلمان و مدارس به دلیل تحقق نیافتن آن چیزی است که امکان دارد خارج از دسترس‌شان و به دلایلی باشد که چندان مرتبط با محدودیت‌های آنها نیست. منطق «هیچ کودکی نباید عقب بماند» و «مسابقه تا اوج» و برنامه‌های مشابه، مسئولیت پایان دادن به شکاف‌ها را بر عهده افرادی می‌گذارد که قدرت و توانایی انجام آن را ندارند و همین امر می‌تواند موجب تضعیف روحیه آموزگاران شود. آموزگاران با معضلات دیگری نیز مواجهند: پیگیری اهداف چندگانه آموزشی در برابر تدریس برای آزمون و دنبال کردن وظایف بزرگ کاری خود در برابر پایبندی به معیارهای

کوچکی که بر اساسشان به آنها پاداش تعلق می‌گیرد. آنها از هر سوی که دست به انتخاب بزنند، بازنده هستند. مضاف بر این، بسیاری از معلمان این روش ایجادشده با فرهنگ آزمون، پاسخگویی و سنجش را به منزله رژیم می‌دانند که استقلال آنها را ربوده و توانایی استفاده از ایده و خلاقیتشان را در طراحی و اجرای برنامه‌های درسی دانش‌آموزان از بین برده است. نتیجه این امر، موج بازنشستگی آموزگاران باتجربه بود و یک حرکت خلاقانه از مدارس دولتی به سوی مدارس خصوصی که با رژیم پاسخگویی معیاری محدود نشده‌اند. [۳۳]

بنابراین تجلیل نفس‌افرادی که بر پاداش مبتنی بر عملکرد آموزشی اندازه‌گیری شده در راستای پایان دادن به شکاف موفقیت پافشاری می‌کنند به بهای نادیده گرفتن کسانی تمام می‌شود که واقعا برای آموزش کودکان می‌کوشند. همه چیزهایی را که می‌توان اندازه‌گیری کرد، نمی‌توان حداقل با اندازه‌گیری بهبود داد.

## فصل ۹

## پزشکی

معیارها در هیچ عرصه‌ای بیشتر از پزشکی رواج ندارند و شاید در هیچ بخشی به اندازه این حوزه، نویددهنده و البته پرخطر نباشند. با این حال معیارها در پزشکی نیز نقش‌های گوناگونی دارند؛ بعضی از آنها حقیقتاً مفیدند و ارزش بعضی دیگر مورد تردید است.

یکی از نقش‌های معیارها، اطلاعاتی و تشخیصی است: روند پیگیری روش‌ها و رویه‌های مختلف و سپس مقایسه خروجی‌ها که امکان تعیین موفق‌ترین روش‌ها و رویه‌ها را می‌دهد، سپس دیگران می‌توانند موفق‌ترین روش‌ها و رویه‌ها را دنبال کنند.

نقش دیگر، معیارهای گزارش شده به صورت عمومی است و هدف آن ایجاد شفافیت برای مشتریان و مصرف‌کنندگان است و زمینه‌ای برای مقایسه و رقابت میان ارائه‌دهندگان ایجاد می‌کند.

نقش دیگر، پرداخت مبتنی بر عملکرد است که در آن پاسخگویی با پاداش‌ها یا مجازات‌های پولی حمایت می‌شود. طرفداران استفاده از معیارها در پزشکی، اغلب درباره این نقش‌های متفاوت در آن واحد بحث می‌کنند.

فشارهای زیاد در دهه‌های اخیر موجب شده است که معیارها نه تنها در جهت افزایش ایمنی و اثربخشی که در زمینه هزینه‌ها نیز مورد استفاده قرار گیرند.

### فشار مالی برای کنترل هزینه‌ها

انگیزه استفاده از معیارها برای کنترل هزینه‌ها ناشی از دلایل مختلف و متعدد است. سالهاست هزینه‌های پزشکی بسیار بیشتر از درآمد ملی در حال افزایش است و این روند افزایشی، دست‌کم تا دهه آینده ادامه خواهد داشت: در سال ۲۰۱۴ بخش سلامت ۱۷/۵ درصد از اقتصاد آمریکا را تشکیل می‌داد و انتظار می‌رود این رقم تا ۲۰۲۵ به ۲۰/۱ درصد برسد. این روند افزایشی دلایل خوبی دارد: هزینه‌های سلامت چیزی است که اقتصاددانان آن را «کالای لوکس» می‌نامند؛ هرچه مردم ثروتمندتر باشند، تمایل بیشتری برای هزینه در این حوزه دارند. واقعیت این است که با افزایش سن جمعیت خردسال، بخش بزرگی از آنها به سنی می‌رسند که دارای بیشترین هزینه‌های پزشکی هستند. به این مورد، دسترسی بیشتر به داروهای تخصصی و رشد سریع قیمت‌های دارو را نیز اضافه کنید. تصویب قانون مراقبت مقرون به صرفه<sup>۱</sup> به معنای این است که درصد بیشتری از هزینه‌های مراقبت‌های بهداشتی و سلامت در ایالات متحده توسط دولت پرداخت می‌شود. با توجه به سهم کل هزینه‌های سلامت که توسط دولت فدرال و دولت ایالتی پرداخت می‌شود و پیش‌بینی‌ها حاکی از آن است که این میزان تا سال ۲۰۲۵ به ۴۷ درصد برسد.<sup>[۱]</sup> افزایش فزاینده مراقبت‌های بهداشتی موجب شده است تا بیمه‌گذاران خصوصی و دولتی (خدمات ملی بهداشت<sup>۲</sup> در انگلیس، مدی‌کر<sup>۳</sup>، مدی‌کد<sup>۴</sup> و اداره کهنه سربازان<sup>۵</sup> در ایالات متحده آمریکا) به پزشکان و بیمارستان‌ها برای کاهش نرخ‌های بازپرداخت و بهبود نتایج فشار

1. The Affordable Care Act

2. The National Health Service

3. Medicare

4. Medicaid

5. The Veterans Administration

بیاورند. هم‌زمان با افزایش فشار برای کنترل هزینه‌ها، فناوری جدید ضبط سوابق الکترونیکی سلامت موجب شده است تا گردآوری داده‌های پزشکی بیشتر در دسترس باشد و موجب ایجاد وسوسه در بهره‌برداری از داده‌ها برای شناسایی و تشخیص مشکلات شود. نتیجه، افزایش حجم گزارش‌های عمومی و پرداخت مبتنی بر عملکرد بود که هر دو مورد به عنوان درمان‌هایی برای بیماری سیستم مراقبت بهداشتی در ایالات متحده آمریکا و خارج از آن شناخته شد و مورد تایید قرار گرفت.

مشکلات واقعی به اندازه کافی هستند: پرداخت‌کنندگان شخص ثالث، چه شرکت‌های بیمه و چه سازمان‌های دولتی مانند مدی‌کد و مدی‌کر نیاز به مدارک معتبری دارند که پزشکان و بیمارستان‌ها خدمات را به شکلی موثر و مقرون به صرفه ارائه می‌دهند اما ثابت شده گاهی این درمان‌ها درست به اندازه بیماری‌هایی که باید معالجه شود، بد است.

### رتبه‌بندی سیستم پزشکی آمریکا

اما پیش از بررسی این درمان‌های ارائه‌شده، بهتر است نگاهی به تاثیرگذارترین معیارهای عملکرد مورد استفاده در توصیف سیستم مراقبت بهداشت و درمان آمریکا بیندازیم؛ معیارهایی که غالباً در حکم مدرکی مبتنی بر لزوم پاسخگویی بیشتر و پرداخت مبتنی بر عملکرد اندازه‌گیری شده ذکر شده‌اند. این معیارها از «گزارش بهداشت جهانی سازمان بهداشت جهانی در سال ۲۰۰۰» به دست آمده‌اند که رتبه ۳۷ را برای سیستم مراقبت‌های بهداشتی ایالات متحده آمریکا را در میان کشورهای جهان تعیین کرده و یادآور شده است که «سخت است این موضوع نادیده گرفته شود که ایالات متحده از نظر هزینه سرانه مراقبت‌های بهداشتی رتبه اول را دارد، اما دارای رتبه ۳۹ در مرگ‌ومیر نوزادان، رتبه ۴۳ در مرگ‌ومیر زنان بزرگسال، رتبه ۴۲ در مرگ‌ومیر مردان بزرگسال و رتبه ۳۶ در امید به زندگی است.» [۲] اسکات دبیلو. اطلس<sup>۱</sup> که پزشک و تحلیلگر سیستم مراقبت‌های بهداشتی و سلامت

1. Scott W. Atlas

بود، این ادعاها را به طور دقیق مورد بررسی قرار داده است و به نظر می‌رسد این موارد بیش از یک مغلطه کوچک باشد.

بیشتر ما تصور می‌کنیم که رتبه‌بندی‌های سازمان بهداشت جهانی، سطح کلی سلامت را اندازه‌گیری می‌کند در حالی که نتایج واقعی سلامت، تنها ۲۵ درصد از میزان رتبه‌بندی را به خود اختصاص داده‌اند. نیمی از امتیازات اعطاشده در راستای عدالت‌خواهی بود: ۲۵ درصد برای «توزیع بهداشت» و ۲۵ درصد دیگر برای «عدالت مالی»؛ جایی که «عدالت» چنین تعریف شده است که همه افراد درصد مشابهی از درآمد خود را برای مراقبت‌های بهداشتی پرداخت کنند. این یعنی که در یک سیستم اگر شما ثروتمندتر باشید، بیشتر نیز برای مراقبت‌های بهداشتی هزینه می‌کنید و آن را عادلانه می‌پندارید. به طور خلاصه این معیار، ایدئولوژیک بود. اما واقعیت این است که این معیار دارای ۳۷ ضمیمه بود که به آن ظاهری قابل اطمینان و بی‌طرف می‌داد. [۳] اما در حقیقت، رتبه‌بندی کلی یک عملکرد همراه‌کننده و فریبنده است.

آمار و ارقام مرگ‌ومیر و امید به زندگی چگونه است؟ ظاهراً این موارد تا حد زیادی متأثر از عوامل خارج از نظام پزشکی است؛ عواملی مرتبط با فرهنگ و سبک زندگی. چاقی منجر به بروز بیماری‌های مزمن و تضعیف‌کننده مانند دیابت نوع دوم و ناراحتی قلبی می‌شود و آمریکایی‌ها به طور متوسط از شهروندان دیگر کشورها چاق‌ترند (اگرچه، تعداد چاق‌ها در برخی کشورها نیز به سرعت در حال افزایش است). کشیدن سیگار نیز می‌تواند در بروز بیماری‌های قلبی، سرطان و دیگر بیماری‌های مزمن نقش داشته باشد و ممکن است حتی پس از دهه‌ها ترک آن باز هم موجب بروز چنین بیماری‌هایی شود. مشخص است که آمریکایی‌ها بنا بر استانداردهای بین‌المللی تا سال‌های ۱۹۸۰ میلادی سیگاری‌های قهاری بودند. نرخ بالای مرگ‌ومیر آمریکایی‌ها با شلیک گلوله، یکی دیگر از عوامل تاسف‌آور است اما تقریباً هیچ ارتباطی با نظام پزشکی ندارد. [۴] علاوه بر این، ایالات متحده از نظر قومی یک کشور ناهمگون است و برخی گروه‌های قومی (مانند آمریکایی‌های

آفریقایی تبار) دارای نرخ بالای مرگ‌ومیر نوزادان هستند و این امر بازتاب عوامل اجتماعی، فرهنگی و احتمالاً ژنتیکی است. [۵] به طور خلاصه، بسیاری از مشکلات سلامت آمریکا تابعی از نظام پزشکی نیست بلکه وابسته به عوامل اجتماعی و فرهنگی خارج از نظام است. هنگامی که صحبت از تشخیص و درمان بیماری است، اطلس یادآور می‌شود که پزشکی آمریکا جزو بهترین‌ها در جهان است. [۶]

در اینجا مانند سایر حوزه‌ها همچون آموزش و امنیت عمومی، بسیاری از مهم‌ترین عواملی که به طور نسبی موجب موفقیت یا شکست می‌شود، خارج از سیستم‌های رسمی قرار دارند که ما تلاش در سنجیدن یا پاسخگو بودن آنها می‌کنیم. ورزش کافی، درست غذا خوردن، دور نگه داشتن سلاح‌های گرم از افراد غیرمسئول و خودداری از کشیدن سیگار، مصرف بیش از حد الکل و مواد مخدر و رابطه جنسی پرخطر؛ اینها مهم‌ترین عوامل موثر در سلامتی و طول عمر هستند. پزشکان و مسئولان سلامت عمومی باید تلاش کنند این موارد را انجام دهند یا بر آنها اثر بگذارند؛ هرچند این الگوهای سبک زندگی خارج از کنترل آنهاست و ما باید این مساله را در ارزیابی نارسایی‌های پزشکی آمریکا در نظر داشته باشیم. حتی اگر به معیارهای هشداردهنده گزارش سازمان جهانی بهداشت با دیده شک و تردید بنگریم، این گزاره باز هم همچنان صحیح است که سیستم بهداشت و سلامت ایالات متحده آمریکا گران و نیازمند بهبود است.

### معیارها در حکم راه حل

شاید محبوب‌ترین روند در سیاست بهداشت و سلامت آمریکا ترویج معیارهای عملکرد، پاسخگویی و شفافیت باشد. تصور می‌شود که عملکرد ارزیابی شده به پزشکان این امکان را می‌دهد تا شیوه‌های بالینی را بهتر ارزیابی و اجرای آنها را پیگیری کنند، به بیمه‌گذاران اجازه می‌دهد تا به موفقیت پاداش دهند و شکست را مجازات کنند و از طریق رتبه‌بندی و گزارش، شفافیت را به شیوه‌هایی ایجاد کنند که به بیماران امکان

انتخاب آگاهانه‌تر در مورد ارائه‌دهندگان خدمات پزشکی می‌دهد. یک از حامیان معیارها، مایکل ای. پورتر<sup>۱</sup> از مدرسه کسب‌وکار هاروارد و طراح «دستور کار ارزش<sup>۲</sup>» است که شامل به‌کارگیری معیارهای مدیریت در پزشکی است. او ادعا می‌کند:

پیشرفت سریع در هر زمینه‌ای نیاز به اندازه‌گیری نتایج دارد که یک اصل آشنا در مدیریت است. گروه‌ها با دنبال کردن پیشرفت در طول زمان و مقایسه عملکرد خود با دیگر گروه‌ها در داخل و خارج سازمان به توسعه و تعالی دست می‌یابند. در حقیقت، اندازه‌گیری دقیق ارزش (نتایج و هزینه‌ها) شاید مهم‌ترین گام در بهبود سلامت و مراقبت‌های بهداشتی باشد. در هر کجا که ما اندازه‌گیری سیستماتیک نتایج را مشاهده کنیم - مهم نیست کدام کشور - متوجه بهبود نتایج خواهیم شد. [۷]

پورتر طرفدار سرسخت گزارش‌دهی نتایج به صورت عمومی بود، زیرا چنین تصور می‌کرد که انگیزه‌ای قدرتمند برای بهبود عملکرد فراهم می‌کند. این تصور به شکل یک نظریه منطقی بود.

### سه روایت از موفقیت

پورتر به کلینیک کلیولند به منزله یک پیشگام رویکرد توصیه‌شده خود اشاره می‌کند. این کلینیک هر سال ۱۴ «کتاب نتایج» را منتشر می‌کند و در آن عملکرد خود را در درمان انواع مختلف و چشمگیر بیماری‌ها به طور مستند نشان می‌دهد. با نگاهی به این اسناد (که به صورت آنلاین در دسترس است) می‌توان نرخ بالای موفقیت را در هر گروه دید. کلینیک کلیولند بیماران بسیاری را از سراسر جهان به خود جلب می‌کند.

یک نمونه قانع‌کننده از فضایل بالقوه معیارهای پزشکی که مایکل پورتر نیز ارائه کرده، از سیستم سلامت جیسینجر<sup>۳</sup> نشأت گرفته است،

1. Michael E. Porter

2. Value Agenda

3. Geisinger Health System

یک سیستم یکپارچه غیرانتفاعی و تحت نظارت پزشک که به حدود ۲/۶ میلیون نفر در پنسیلوانیا خدمت می‌کند که بسیاری از آنها متعلق به قشر فقیرند. جیسینجر به دلایل قانع‌کننده، ویترونی برای متری نشان دادن سیستم سلامت و مراقبت بهداشتی در ایالات متحده آمریکا است. [۸] جیسینجر یکی از پیشگامان استفاده از سوابق الکترونیکی پزشکی بود که در سال ۱۹۹۵ بیش از ۱۰۰ میلیون دلار را صرف سیستم ضبط الکترونیکی سوابق خود کرد و همین امر به پزشکان انگیزه داد تا بیماران خود را در یک پورتال آنلاین ثبت نام کنند. این سیستم امکان انتقال اطلاعات را به ارائه‌دهندگان سیستم و نظارت بر عملکرد واحدها از جمله پزشکان می‌داد. این سیستم، مددجویان پرستار را برای بیماران پرخطر استخدام می‌کند تا به بیماران درباره وضعیت‌شان آموزش دهند، بر آنها نظارت داشته باشند، برنامه‌ها و داروهای‌شان را بررسی کنند و پس از تشخیص یا حتی درمان با آنها قرار ملاقات بگذارند. دو بیماری شایع و پرهزینه در نظام سلامت آمریکا، دیابت و بیماری‌های قلبی است. در سیستم جیسینجر، بیماران در چنین شرایطی توسط یک تیم یکپارچه متشکل از پزشکان و دستیاران‌شان، داروسازان، متخصصان تغذیه و ... تحت درمان قرار می‌گیرند. جیسینجر به جای ارائه درمان با مجموعه‌ای که از هم جدا و دارای تماس حداقلی با یکدیگرند، رویکرد جامع‌تری را به کار می‌گیرد. حدود ۲۰ درصد از پرداختی پزشکان با اهداف مربوط به کاهش هزینه‌ها، بهبود کیفیت مراقبت و رضایت بیمار گره خورده است در حالی که ۸۰ درصد دیگر بر اساس هزینه پرداخت‌شده برای خدمات است. جیسینجر از طریق برنامه‌های نوآورانه خود موفق به کاهش هزینه‌ها و بهبود بیماران شد. یکی از کاربردهای بدون ابهام و موفقیت‌آمیز معیارها در پزشکی استفاده از عملکرد اندازه‌گیری‌شده برای کاهش عفونت‌های بیمارستانی در «خطوط مرکزی»<sup>۱</sup> بدن است. خطوط مرکزی لوله‌های منعطف سوند هستند که به عنوان مجرای برای غذا، دارو یا مواد مغذی در رگ (ورید) بزرگ در نواحی گردن یا قفسه

سینه جای داده می‌شوند. خطوط مرکزی از مهم‌ترین عناصر پزشکی بیمارستانی مدرن به شمار می‌روند و تا همین چند سال اخیر یکی از مهم‌ترین مواردی بود که مشکلات زیادی را در پی داشت. دلیلش هم این است که سوندها می‌توانند راهی برای ایجاد عفونت باشند، آنها در بدترین حالت گشوده هستند و در بهترین موارد نیز هزینه بالایی دارند. تخمین زده می‌شد که در سال ۲۰۰۱ در ایالات متحده حدود ۸۲ هزار عفونت خونی ناشی از سوند وجود داشته است. هزینه به ازای هر عفونت بین ۱۲ تا ۵۶ هزار دلار متغیر بود و در حدود ۳۲ هزار نفر نیز جان باختند.[۹]

از آن زمان، به لطف تلاش‌های پیتر جی. پرونوست<sup>۱</sup>، متخصص مراقبت‌های ویژه بیمارستان جان هاپکینز در بالتیمور میزان عفونت‌ها به طور چشمگیری کاهش یافته است. او با همکاران خود برنامه‌ای مبتنی بر فهرستی شامل پنج فرآیند استاندارد و در عین حال ساده تهیه کرد که در مجموع احتمال عفونت در خطوط مرکزی را کاهش می‌داد. پرونوست پس از اجرای برنامه خود در جان هاپکینز بر اجرای آن تحت عنوان «پروژه آی‌سی‌یو کی‌استون میشیگان»<sup>۲</sup> در بیمارستانی در میشیگان نظارت داشت. از آن زمان، پروژه‌های مشابهی در سراسر ایالات متحده آمریکا، انگلستان و اسپانیا اجرا شده است. نتایج نیز بسیار چشمگیر بوده است: عفونت‌های جریان خون حدود ۶۶ درصد کاهش یافته و موجب نجات جان هزاران انسان و حفظ میلیون‌ها دلار شده است.

پروژه کی‌استون شامل جمع‌آوری داده‌های ماهانه درباره میزان عفونت است که به روسای واحدهای مراقبت‌های ویژه و مقامات ارشد بیمارستان گزارش می‌شود. نتایج با گروه‌های کاری بزرگ‌تر و با نگاهی به یادگیری از اشتباهات، قابل بحث است. این مورد، نمونه‌ای از معیارهای تشخیصی است. این پروژه، داده‌هایی را ارائه می‌دهد که توسط یک فرد شاغل (پزشک) یا در داخل یک نهاد (بیمارستان) استفاده یا در میان شاغلان و نهادها به اشتراک گذاشته می‌شود تا مشخص شود چه چیزی درست

1. Peter J. Pronovost

2. Michigan Keystone ICU Project

کار می‌کند و برعکس که در نهایت این اطلاعات برای بهبود عملکرد مورد استفاده قرار می‌گیرد.

پروژه‌کی‌استون شامل استفاده گسترده از معیارهای تشخیصی و برخی انگیزه‌های روانی مانند فشار همتایان است. خود پروروست موفقیت خود را به حساب این واقعیت گذاشت که این پروژه بواسطه مجامع بالینی انجام می‌شود؛ کار در راستای اهداف حرفه‌ای مشترک و درمان عفونت‌های خطوط مرکزی به منزله حل یک معضل اجتماعی. دیدن میزان عفونت در مقایسه با سایر بیمارستان‌ها موجب فشار همتایان می‌شود و موجب تلاش بیشتر برای رسیدن به میزان موفقیت همتایان یا پیشی گرفتن از آنها خواهد شد.

### چه نتیجه‌ای باید از این موفقیت‌ها بگیریم؟

کلینیک کلیولند، جیسینجر و پروژه‌کی‌استون دائماً در حکم مدرکی مستدل از کارایی عملکرد اندازه‌گیری‌شده ذکر می‌شود. با این حال زمانی که عمیق به موضوع نگاه کنیم متوجه خواهیم شد که معیارها به دلیل قرار گرفتن در یک فرهنگ نهادی بزرگ‌تر حائز اهمیت است. آیا موفقیت کلینیک کلیولند تابعی از این واقعیت است که این کلینیک نتایج خود را منتشر می‌کند؟ یا این کلینیک واقعا مشتاق است به این دلیل که نتایجش موثر و چشمگیرند، آنها را منتشر کند؟ کلینیک کلیولند پیش از ظهور معیارهای عملکرد یکی از بزرگ‌ترین موسسات عالی پزشکی در جهان بود و در حقیقت جایگاه خود را در عصر معیارهای عملکرد نیز حفظ کرده اما این نتیجه‌گیری که یک رابطه سببی میان کیفیت کلینیک و انتشار عملکرد معیارها وجود دارد، یک مغالطه است. این موفقیت شاید بیشتر وابسته به شرایط محلی باشد - روش‌هایی که فرهنگ سازمانی کلیولند از معیارها استفاده می‌کند - تا کیفیت خود اندازه‌گیری. [۱۰]

معیارها در برنامه جیسینجر به دلیل روشی که در یک سیستم بزرگ‌تر تعبیه شده، موثر است. اساساً، تعیین معیارهای اندازه‌گیری و

ارزیابی عملکرد توسط تیم‌هایی متشکل از پزشکان و مدیران (سرپرستان) انجام می‌شود. بنابراین معیارهای عملکرد توسط مدیرانی تهی از دانش لازم، تحمیل و ارزیابی نمی‌شود بلکه بر اساس همکاری و بازبینی همتایان ایجاد می‌شود. جیسینجر همچنین از معیارهای خود برای بهبود مداوم عملکرد در مراقبت‌های سرپایی در شرایط مختلف استفاده می‌کند. در اینجا به صحبت‌های گلن دی. استیل<sup>۱</sup> اشاره می‌کنیم؛ پزشکی که مدیرکل نظرت بر تحول سیستم جیسینجر بود و درباره موفقیت‌های‌شان چنین گفت: «مسیرهای مراقبت جدید ما بسیار موثر و کارآمد است زیرا توسط پزشکان هدایت و با بازخوردهای داده‌محور در زمان واقعی تهیه می‌شوند و در درجه اول بر بهبود کیفیت مراقبت از بیمار تمرکز دارند که اساساً موجب ایجاد انگیزه برای تغییر رفتار پزشکان ما است.» واقعیت بسیار مهم این بود که «هر زن و مردی که در خطوط خدمات کار می‌کند خود تصمیم می‌گیرد که مراحل مراقبت را تغییر دهد. درگیر شدن آنها در فرآیند تصمیم‌گیری به طور مستقیم، هم امنیت آنها را تضمین و هم احتمال موفقیت‌شان را بیشتر می‌کند.» آنچه می‌توانیم از مثال جیسینجر بیاموزیم، اهمیت داشتن اراده‌دهندگان برای ترویج و نظارت بر اندازه‌گیری عملکرد است. این واقعیت که اندازه‌گیری‌ها مطابق با احساس رسالت حرفه‌ای آنها بوده، بسیار اهمیت دارد.

پیتر پرونوست که پیشگام کاهش عفونت‌های خطوط مرکزی بود، بر این باور است که «پروژه کی‌استون آی‌سی‌یو، پتانسل تلاش‌های داوطلبانه را نشان می‌دهد که متکی به انگیزه ذاتی بواسطه هنجارهای همتایان و حرفه‌ای بودن است.» او موافق تکمیل این درخواست‌ها از طریق گزارش‌نویسی عمومی و مشوق‌های پولی نیست اما تفسیر او این است که بهبود در نتایج پزشکی در درجه اول با «تغییر در باور پزشکان انجام می‌شود، باور به اینکه میزان عفونت گریزناپذیر نیست و قابل کنترل است؛ آن هم به روشی که برای اخلاق حرفه‌ای پزشکان و

1. Glenn D. Steele

پرستاران جذاب باشد.»

با این حال، نتیجه‌ای که توسط مراکز دولتی ایالات متحده برای مدی‌کر و مدی‌کد گرفته شده این است که قرار بود این دو مرکز درباره میزان عفونت در سال ۲۰۱۱ گزارش دهند و یک سال بعد بیمارستان‌های با میزان بیشتر عفونت را با جلوگیری از پرداخت به آنها، جریمه کنند. این امر، ساختاری از انگیزه‌های بسیار متفاوت از موفقیت‌های نهادی که تاکنون بررسی کرده‌ایم ایجاد کرده که بیشتر به انگیزه‌های ذاتی و درونی وابسته است تا بیرونی.

**تصویر بزرگ‌تر: معیارها، پرداخت مبتنی بر عملکرد، رتبه‌بندی و گزارش**  
 هنگامی که در سوابق عملکرد معیارها در حوزه پزشکی به طور عمیق‌تری کنکاش کنیم، به نظر می‌رسد موفقیت‌های کلینیک کلیولند، برنامه جیسینجر و پروژه کی‌استون جزء استثنائات است. بیشتر متخصصانی که درباره معیارهای پزشکی می‌نویسند، علاقه بسیاری به اثربخش بودن اندازه‌گیری عملکرد دارند اما مشاغل آنها ارتباطی با کارایی جمع‌آوری و تحلیل داده‌ها ندارد. از این رو، بسیاری از مطالعات نشان‌دهنده نبود یا میزان ناچیزی از کارایی در انتشار پاسخگویی معیارهاست که باید آن را غلبه مدرک بر علاقه دانست. خواهیم دید که مجلات مراقبت‌های پزشکی و ادبیات دانشگاهی از چنین مطالعاتی اشیاع است. آنها برای اطمینان بیشتر به دنبال داده‌ها و مطالعات بیشتر و معیارهای دقیق‌تر هستند تا اینکه بی‌مهابا اعلام کنند که معیارها بی‌فایده بوده‌اند. [۱۱]

اما واقعیت این است که مطالعات شکست‌خورده اغلب توسط افرادی انجام می‌شود که با عملکرد اندازه‌گیری شده نه تنها مخالفتی ندارند که حتی برای‌شان بسیار مهم است. [۱۲]

استدلال برای پاسخگویی و شفافیت مبتنی بر این فرض است که انتشار عمومی معیارهای موفقیت و شکست بر رفتار بیماران، متخصصان و سازمان‌ها اثرگذار خواهد بود. بیماران با مقایسه هزینه‌های مراقبت و میزان نسبی موفقیت مانند مصرف‌کنندگان عمل می‌کنند. پزشکان

متخصصانی را که نمرات عملکرد بالایی دارند به بیماران توصیه می‌کنند. بیمه‌گران از بیمارستان‌ها و ارائه‌دهندگان‌گی که بهترین خدمات را با کمترین قیمت ارائه می‌دهند، بازدید می‌کنند. پزشکان و بیمارستان‌ها برای بهبود نمرات خود تحت فشار خواهند بود تا شهرت و درآمدشان آسیب نبینند. [۱۳]

برای بررسی درستی این نظریه در واقعیت، گروهی از متخصصان پژوهشگاه علمی کیفیت بهداشت و درمان (بهداشت و درمان آی کیو)<sup>۱</sup> در مرکز پزشکی دانشگاه رادبود در نیمخن<sup>۲</sup> هلند شواهد موجود را بررسی کردند تا ببینند اطلاعات در دسترس در مورد پژوهش‌های مرتبط با انواع موضوعات مختلف سلامت چگونه بر رفتار ارائه‌دهنده و مصرف‌کننده (بیمار) و نتایج اثرگذار بوده است. آنها شامل مطالعاتی کنترل‌شده‌ای هستند که رفتار را قبل و بعد از معرفی معیارهای پزشکی در دسترس برای طیف گسترده‌ای از بیماری‌ها مانند حملات قلبی مقایسه می‌کنند. کارشناسان هلندی دریافتند که در برخی موارد، بیمارستان واقعا در روند کار خود پیشرفت‌هایی دارد اما برخلاف پیش‌بینی طرفداران پاسخگویی، این امر هیچ تاثیر ماندگاری بر نتایج بیماران نداشت.

این امر می‌تواند محصول ارتباط میان پژوهش‌های پزشکی و عمل پزشکی باشد. جمعیتی که روی آنها تحقیقات پزشکی انجام شده است متفاوت از جمعیت واقعی هستند که در بیمارستان‌ها و توسط پزشکان تحت درمان قرار گرفته‌اند. مداخلات پزشکی احتمالی (مانند کنترل قند خون برای پیشگیری از ابتلا به دیابت) روی گروه‌های نسبتا کوچکی از بیماران انجام شده است و برای جدا کردن اثرات مداخله، چنین مطالعاتی به طور عمدی بیماران دارای مشکلات متعدد پزشکی را از گروه‌های آزمایش حذف می‌کند. اما در دنیای واقعی، بیماران معمولا دارای چندین مشکل پزشکی به طور همزمان هستند (هم‌آیندی اختلالات<sup>۳</sup>)، از این رو اثر مداخله آزمایش‌شده از بین می‌رود. همین موضوع می‌تواند توضیح

1. Scientific Institute for Quality of Healthcare (IQ Healthcare)

2. Radboud University Nijmegen Medical Centre

3. Comorbidity

دهد که چرا صرفاً پیروی از مراحل توصیه‌شده لزوماً به بهبود نتایج نمی‌انجامد. [۱۴] بنا به گفته کارشناسان هلندی، انتشار معیارها بر رفتار بیمار در انتخاب ارائه‌دهنده خدمات پزشکی یا بیمارستان اثر نداشته است. نتیجه‌گیری آنها: «مجموعه کوچکی از شواهد در دسترس نشان می‌دهد که هیچ مدرک ثابتی مبنی بر اینکه انتشار داده‌های عملکرد، موجب تغییر رفتار مصرف‌کننده می‌شود یا روند مراقبت را بهبود می‌دهد، وجود ندارد.» [۱۵]

یکی دیگر از کاربردهای برجسته معیارها، برنامه‌های پرداخت مبتنی بر عملکرد است. در اینجا ساختار تشویقی ساده است: پزشکان بخش مهمی از پاداش خود را برای رسیدن به برخی اهداف اندازه‌گیری‌شده مانند دنبال کردن مراحل توصیه‌شده (فهرست‌ها) یا کاهش هزینه‌ها یا بهبود نتایج دریافت می‌کنند.

خدمات ملی بهداشت<sup>۱</sup> در بریتانیا آغاز به پذیرش پرداخت مبتنی بر عملکرد، به منزله یک عامل مهم طبقه‌بندی پرداخت به پزشکان مراقبت‌های اولیه در اواسط دهه ۱۹۹۰ کرد و این برنامه توسط دولت تونی بلر تمدید شد. در ایالات متحده آمریکا، برنامه‌های بهداشت و سلامت خصوصی و اجراکنندگان آن مانند دولت‌های ایالتی به طور فزاینده‌ای برنامه‌های پرداخت مبتنی بر عملکرد را پذیرفتند. مقررات پرداخت مبتنی بر عملکرد بخش مهمی از پاداش پزشکی است که پزشکان آن را از مدی‌کر به عنوان بخشی از قانون مراقبت مقرون به صرفه سال ۲۰۱۰ دریافت می‌کنند. [۱۶] مدیران مدی‌کر تلاش کرده‌اند تا پاداش انواع نتایج اندازه‌گیری‌شده، از جمله نتایج جراحی را به عنوان معیار میزان زنده ماندن تا سی روز پس از جراحی در نظر بگیرند.

شکل برجسته دیگر معیارهای پزشکی، رتبه‌بندی عمومی پزشکان و بیمارستان‌ها در قالب «برگه‌های گزارش پزشکی<sup>۲</sup>» است. ایالت نیویورک پیشگام انتشار چنین داده‌هایی بود؛ در انگلستان وزارت بهداشت در سال

1. National Health Service (NHS)

2. Medical Report Cards

۲۰۰۱ شروع به «ستاره دادن»<sup>۱</sup> به سازمان‌های بهداشت عمومی کرد و انگلستان اخیراً به تنها کشوری بدل شد که انتشار «نتایج داده‌ها» را در ۹ جراحی تخصصی برای جراحان منتشر کرده است. در سال ۲۰۱۵ سازمان گزارشگران آمریکایی پروپابلیکا<sup>۲</sup> مشکلات حدود ۱۷ هزار جراح در سراسر ایالات متحده را منتشر کرد. [۱۷] همچنین برگه‌های گزارش و رتبه‌بندی‌ها توسط «کمسیون مشترک» غیرانتفاعی در مورد اعتبار پزشکی و توسط رتبه‌بندی‌های سودآور<sup>۳</sup> خصوصی مانند وبسایت‌های هِلث‌گریڈز<sup>۴</sup> و یو‌اِس نیوز<sup>۵</sup> و وُردِر پورِت (گزارش جهانی)<sup>۶</sup> منتشر می‌شود. تصور همه این گروه‌ها این است که پزشکان و بیمارستان‌ها در راستای ارتقای اعتبار خود برای امنیت و کارایی و در نهایت افزایش سهم بازار از جمعیت بیماران بالقوه، دارای انگیزه داشتن عملکرد بهتر خواهند بود. برای بیمارستان‌ها، اهمیت این رتبه‌بندی‌ها در راستای بهبود وضعیت و «مدیریت نامانام»<sup>۷</sup> است. [۱۸]

هم‌اکنون ادبیات علمی اجتماعی بزرگی درباره تاثیر پرداخت مبتنی بر عملکرد و معیارهای عملکرد عمومی در ایالات متحده آمریکا، بریتانیا و جاهای دیگر وجود دارد. شگفت‌انگیز این است که اغلب این تکنیک‌ها -که مطابق نظریه اقتصادی کاملاً موثرند- هیچ تاثیر قابل توجهی بر نتایج ندارند. [۱۹] مثلاً مطالعه جدیدی در سالنامه پزشکی داخلی به بررسی سرنوشت بیمارانِ مِدی‌کِر از سال‌های آغاز گزارش عمومی میزان مرگ‌ومیر بیمارستان در سال ۲۰۰۹ می‌پردازد. به گفته نویسندگان: «ما متوجه شدیم که انتشار عمومی آمار مرگ‌ومیر تاثیری بر نتایج بهبود بیماران نداشته است. ما نگاهی به هر زیرگروه انداختیم. ما حتی به بررسی افرادی که عملکردشان بد معرفی شده بود پرداختیم تا ببینیم آیا

1. Star Ratings
2. ProPublica
3. Profit-making
4. Healthgrades
5. US News
6. World Report
7. Brand Management

عملکردشان با سرعت بیشتری بهبود می‌یابد. خیر؛ اینطور نبود. در حقیقت اگر می‌خواهید به داده‌ها وفادار باشید نتیجه می‌گیرید که گزارش‌های عمومی سرعت بهبود بیماران را کاهش می‌دهد.» [۲۰] مهم‌تر از همه، مشکل این است که طبق آنچه متخصصان، معیارهای مشکوک می‌نامند، بسیاری از رتبه‌بندی‌های عمومی مانند برگه گزارش جراحی پروپابلیکا به جای آنکه واقعا روشنگرانه باشد، گمراه‌کننده است. [۲۱]

گزارش جدید دیگری این بار از طرف شرکت رندا<sup>۱</sup> به نتیجه‌گیری مشابهی رسیده است. این مطالعه یادآور این نکته شد که مانند بیشتر مطالعات در زمینه پرداخت مبتنی بر عملکرد، روند کار و نتایج بیمار در این فرآیند مورد بررسی قرار می‌گیرد، نه نتیجه نهایی که آیا بیمار بهبود یافته است یا خیر. بر اساس این پژوهش: «به طور کلی، مطالعات با روش‌شناسی‌های قوی‌تر به‌گونه‌ای طراحی شده‌اند که احتمال بسیار کمتری وجود دارد تا پیشرفت‌های چشمگیر مرتبط با برنامه‌های پرداخت مبتنی بر عملکرد را شناسایی کنند. شناسایی تأثیرات‌شان نیز نسبتاً کم است.» [۲۲] این یافته جدیدی نبود؛ دانشمندان علوم اجتماعی نیز که در سال ۱۹۹۰ طرح‌های پرداخت مبتنی بر عملکرد را در بخش دولتی مطالعه می‌کردند به این نتیجه رسیدند که این طرح‌ها ناکارآمد است. با این حال، چنین طرح‌هایی همچنان معرفی می‌شوند: پیروزی امید بر تجربه، یا مشاورانی که بر همان درمان‌های همیشگی قدیمی تأکید می‌کنند. [۲۳]

معیارهای استفاده‌شده برای رتبه‌بندی عمومی و پرداخت مبتنی بر عملکرد، اغلب به روش‌هایی ناخواسته و خلاف انتظار بر نتایج اثر می‌گذارد. این معیارها -چه مولد و چه غیرمولد- معمولا شامل هزینه‌های کلانی است که به‌ندرت توسط طرفداران پرداخت مبتنی بر عملکرد یا معیارهای شفافیت در نظر گرفته می‌شود. در میان مشکلات پرداخت مبتنی بر عملکرد و رتبه‌بندی عمومی، انحراف هدف نیز وجود دارد. همان‌طور که در یک گزارش انگلیسی

1. Rand Corporation

آمده است «برنامه‌های پرداخت مبتنی بر عملکرد تنها به آنچه که اندازه‌گیری و به آن اعتبار داده می‌شود پاداش می‌دهد؛ محدودیتی که می‌تواند به مراقبت کمتر بینجامد و موجب تمرکز نامتناسب پزشک بر آنچه که اندازه‌گیری می‌شود و نه آنچه که مهم است، شود.» برنامه پرداخت مبتنی بر عملکرد بریتانیا منجر به کیفیت پایین‌تر مراقبت پزشکی برای افرادی شد که بخشی از برنامه نبودند. به طور خلاصه، این برنامه منجر به «درمان برای آزمایش» شد. ارائه معیارهای قابل اطمینان برای اندازه‌گیری درمان بسیاری از بیماران از جمله بیماران نحیف و مسن که از بیماری‌های متعدد و مزمن رنج می‌برند، به‌سادگی غیرممکن است. [۲۴]

برگه‌های گزارش پزشکی نیز به همین اندازه مشکلات را حل می‌کند. پدیده خطرگریزی<sup>۱</sup> را در نظر بگیرید؛ مطالعات بسیاری نشان داده‌اند که جراحان قلب پس از معرفی معیارهای در دسترس عموم، تمایل بسیار کمتری برای جراحی بیمارانی به‌شدت نیازمند جراحی دارند. مثلاً در ایالت نیویورک، برگه‌های گزارش جراحان حاوی اطلاعاتی درباره میزان مرگ‌ومیر پس از عمل جراحی بای‌پس عروق کرونر<sup>۲</sup> است؛ یعنی اینکه چند درصد از بیماران تحت این عمل تا سی روز پس از عمل زنده می‌مانند. پس از آغاز استفاده از معیارها، میزان مرگ‌ومیر واقعاً کاهش یافت که به نظر می‌رسد پیشرفتی مثبت باشد اما تنها آن دسته از بیمارانی که تحت عمل جراحی قرار گرفته بودند، شامل معیارها می‌شدند. بیمارانی که پزشک به دلیل پرخطر بودن آنها از عمل جراحی‌شان سرباز زد - و بنابراین می‌توانستند موجب کاهش امتیاز جراح شوند - شامل این معیارها نیستند. برخی از این بیماران به کلینیک کلیونند ارجاع داده شدند و از این رو نتایج آنها در معیارهای نیویورک نشان داده نشده است. در نتیجه «غرض‌ورزی در انتخاب مورد» (یا همان سرشیرگرفتن) برخی از بیماران به‌سادگی تحت عمل جراحی قرار نگرفتند. همچنین روشن نیست که بهبود نتایج پس

1. Risk-aversion.

2. Coronary Bypass Surgery

از عمل جراحی در ایالت نیویورک، نتیجه انتشار معیارها بوده است یا نه. مشخص شد که پیشرفت مشابهی در ایالت همسایه ماساچوست نیز اتفاق افتاده است؛ جایی که هیچ گزارشی از داده‌ها به صورت عمومی منتشر نشده بود. [۲۵]

پدیده خطرگریزی یعنی برخی بیمارانی که جان‌شان با یک عمل جراحی پرخطر حفظ می‌شود، به‌سادگی هرگز تحت آن عمل قرار نمی‌گیرند. اما مشکل مقابل این مورد نیز وجود دارد؛ مراقبت بیش از حد برای رسیدن به اهداف معیارها. شاید بیمارانی را که عمل جراحی‌شان موفق نبوده به مدت سی روز در راستای بهبود آمار مرگ‌ومیر بیمارستان زنده نگه دارند؛ فرآیندی طولانی، پرهزینه و غیرانسانی. [۲۶]

مطمئناً برخی مزایا نیز در انتشار عمومی معیارهای موفقیت جراحان و میزان مرگ‌ومیر بیمارستان‌ها وجود دارد. انتشار داده‌ها می‌تواند افرادی را که عملکرد ضعیفی دارند و شاید بعدها موجب توقف این فرآیند شوند، شناسایی کند؛ مثلاً در مورد جراحان. در این حرفه فرآیند غربال‌گری ارزشمندتری وجود دارد که پزشکان تمایلی به عزل اعضای بی‌کفایت این صنف ندارند. در مورد بیمارستان‌ها، شاغلان سطوح پایین‌تر می‌توانند برای بهبود عملکرد اندازه‌گیری‌شده اقداماتی انجام دهند. اما در اینجا مانند بسیاری از معیارهای عملکرد، گرایش به این سمت است که با انجام ساده‌ترین کارها، انتظار نتایج فوق‌العاده را داشت. به این معنا که مزایای فوری برای کشف داده‌های خارج از محدوده نسبت به عملکرد ضعیف وجود داشته باشد. [۲۷] مشکل اینجاست که معیارها از همه افراد جمع‌آوری می‌شوند و در برخی موارد هزینه‌های جانبی از مزایای جانبی بیشترند. در گزارش اخیر یکی از نهادهای پزشکی، پرهزینه و سنگین بودن پیگیری هر چه بیشتر معیارهای پزشکی مشهود است. [۲۸] در مراکز پزشکی بزرگ، هزینه گزارش اقدامات باکیفیت برای قانون‌گذاران دولتی و بیمه‌گذاران یک درصد درآمد خالص بود. هزینه‌های اداری برای اندازه‌گیری و فعالیت‌های مرتبط سالانه حدود ۱۹۰ میلیارد دلار تخمین زده می‌شود. همچنین هزینه‌های غیرقابل اندازه‌گیری وارد کردن

داده‌ها به سیستم‌های دولتی گزارش کیفیت بیماران نیز وجود دارد. برای اقدامات پزشکی بزرگ‌تر باید به شرکت‌های خارجی هزینه وارد کردن داده‌ها پرداخت شود؛ در اقدامات کوچک‌تر گاهی این کار توسط خود پزشکان انجام می‌شود. علاوه بر این، هزینه‌های قابل توجه جمع‌آوری و وارد کردن و پردازش این سونامی داده، هزینه‌های غیرقابل محاسبه دیگری نیز هستند؛ مثلاً زمانی که پزشکان و دستیاران‌شان صرف‌و‌اورد کردن داده‌ها می‌کنند. جدا از اینکه زمان صرف‌شده تا حد زیادی غیر قابل محاسبه و جبران است، معمولاً زمانی که هزینه‌های پزشکی مورد بحث قرار می‌گیرد به این هزینه‌ها توجهی نمی‌شود. [۲۹] مطالعه جالب یک موسسه پزشکی گزارش می‌دهد: «رشد سریع علاقه، حمایت و ظرفیت در تلاش برای اندازه‌گیری اهداف مختلف - شامل بهبود و ارزیابی عملکرد، گزارش عمومی و تامین‌کننده مالی و ابتکارات پیشرفت داخلی - اثربخشی این تلاش‌ها را کاهش داده است.»

دونالد ام. برویک<sup>۱</sup> یک مدافع سرسخت پیشرفت از طریق اندازه‌گیری است و به عنوان مدیر و سرپرست مراکز مدی‌کر و مدی‌کد در سال‌های ۲۰۱۰ تا ۲۰۱۱ خدمت کرده است. الزامات گزارش‌نویسی چنان سنگین و زائد شده است که دکتر برویک اخیراً اعلام کرد: «ما باید جلوی اندازه‌گیری بیش از حد را بگیریم ... من به کاهش ۵۰ درصدی کلیه معیارهایی که در حال حاضر استفاده می‌شود، رای می‌دهم.» [۳۰]

به این موارد، هزینه‌های روانی درمان پزشکی را نیز اضافه کنید که به نظر می‌رسد یک سرمایه‌گذاری سودآور است. برویک این موضوع را به طور شگفت‌انگیزی در مقاله خود تحت عنوان «مسمومیت پرداخت مبتنی بر عملکرد» مورد بررسی قرار داده است:

«پرداخت مبتنی بر عملکرد، انگیزه ذاتی را کاهش می‌دهد. بسیاری از وظایف، به‌ویژه در بخش درمان و سلامت به طور بالقوه و ذاتی رضایت‌بخش است. تسکین درد، پاسخ به پرسش‌ها، انجام فعالیت‌های دستی، داشتن اعتماد به نفس، کار با تیم حرفه‌ای، حل معما و تجربه

1. Donald M. Berwick

کردن نقش یک مقام قابل اعتماد؛ هیچ یک از اینها روش‌های بدی برای گذراندن روز کاری یک فرد نیست. افتخار و لذتی در کار مراقبت وجود دارد و در میان دیگر انگیزه‌های متعددی قرار می‌گیرد که منجر به «عملکرد» متخصصان حوزه سلامت و بهداشت می‌شود. در مباحث تلخ مرتبط با جبران، هزینه و پرداخت که امروزه، زمان مدیران حوزه سلامت و بهداشت را می‌گیرد، نادیده گرفتن و شک کردن به این واقعیت که پاداش‌های غیرمالی و ذاتی در این حوزه از اهمیت زیادی برخوردارند، بسیار ساده است. متأسفانه غفلت از رضایت ذاتی در کار می‌تواند به طور ناخواسته آن را کاهش دهد.» [۳۱]

مقاله برویک بیش از دو دهه قبل منتشر شده و به نظر می‌رسد که هیچ تاثیری نداشته است. موج جزرومدی پرداخت مبتنی بر عملکرد همچنان در حال بالا آمدن است.

### یک مورد آزمایش شده: کاهش بازپذیرش‌ها

بیشترین اندازه‌گیری‌های استفاده‌شده، معیارهای مدی‌کر در راستای پذیرش مجدد و برنامه‌ریزی‌نشده بیمارانی هستند که پس از سی روز از بیمارستان‌ها ترخیص شده‌اند و خود، نشان‌دهنده مشکلات معیارها است. پذیرش در بیمارستان‌ها، پرهزینه است و یکی از انگیزه‌های این کار، کاهش هزینه‌هاست. همچنین تصور می‌شود بازپذیرش بیمار نتیجه مراقبت نکردن از اوست، از این رو کاهش تعداد بازپذیرش‌ها نشانه‌ای از بهبود مراقبت خواهد بود. در سال ۲۰۰۹ مدی‌کر شروع به گزارش عمومی به عنوان یک معیار شفافیت از بیمارستان‌هایی کرد که پس از گذشت سی روز از پذیرش مجدد بیماران با مراقبت دقیق آنها را ترخیص کرده بودند. سی روز بستری مجدد بیمارانی را تحت پوشش قرار می‌دهد که برای شرایط سخت و مهم پزشکی مورد درمان قرار گرفته‌اند: سکته قلبی، نارسایی قلبی، سکته مغزی، ذات‌الریه، بیماری انسداد ریوی مزمن، بای‌پس عروق کرنر، دو عمل جراحی تعویض مفصل ران یا زانو (معیارهای منتشرشده مدی‌کر را در بیمارستان و وبسایت با

یکدیگر مقایسه کنید). سپس در سال ۲۰۱۲، مدی کر با اعمال مجازات برای بیمارستان‌هایی با میانگین نرخ بالاتر از انتشار گزارش‌های عمومی به سوی پرداخت مبتنی بر عملکرد رفت. [۳۲] گزارش عمومی از عملکرد و مجازات پولی برای شکست به عنوان یک انگیزه برای بیمارستان‌هاست در راستای اقداماتی برای محدود کردن بازپذیرش و از آنجا که هزینه‌های پذیرش بالاست، این امر در راستای کاهش این هزینه‌ها نیز انجام می‌شود. بیمارستان‌ها آغاز به اقدامات دیگری برای اطمینان از بازنگشتن بیماران پس از ترخیص کردند؛ این اقدامات شامل هماهنگی بهتر با ارائه‌دهندگان مراقبت‌های اولیه و تلاش برای حصول اطمینان از دسترسی بیمار به داروهای تجویز شده است. جریمه‌هایی که برای بیمارستان‌هایی با عملکرد ضعیف بریده می‌شود با قصد ایجاد انگیزه برای ارائه مراقبت بهتر از بیماران است تا آنها مجبور نشوند دوباره به بیمارستان بازگردند. بازپذیرش در بیمارستان‌ها واقعا کاهش یافته است؛ یک موفقیت چشمگیر برای معیارهای عملکرد، اما این موفقیت چقدر واقعیست؟

کاهش میزان بازپذیرش‌های گزارش شده تا اندازه‌ای مربوط به بازی سیستمی بود: به جای اینکه بیماران بازگشتی به طور رسمی پذیرش شوند، بیمارستان‌ها آنها را در «وضعیت تحت نظر» قرار دادند؛ وضعیتی که بیماران برای مدت کوتاهی (نهایتاً تا چند روز) در بیمارستان می‌مانند و به جای «پذیرش» بیمار در حسب حال و صورت حساب آنها خدمات سرپایی درج می‌شود. از سوی دیگر، بسیاری از بیماران بازگشتی نیز در اتاق اورژانس تحت درمان قرار گرفتند. بین سال‌های ۲۰۰۶ تا ۲۰۱۳ چنین وضعیتی برای بیماران مدی کر ۹۶ درصد افزایش یافته است. این یعنی نیمی از افت بازپذیرش‌ها به دلیل بیمارانی بوده است که در حقیقت به بیمارستان بازگشته‌اند اما به عنوان بیمار سرپایی درمان شده‌اند (در راستای پیچیده‌تر شدن امور، تجزیه و تحلیل بعدی نشان داد که بیمارستان‌هایی که بازپذیرش خود را کاهش داده‌اند آنهایی نیستند که بیماران در وضعیت تحت نظر آنها افزایش یافته باشد).

بنابراین معیارهای بازپذیرش بهبود یافته‌اند اما این امر لزوماً به معنای افزایش کیفیت مراقبت از بیمار نیست. همه بیمارستان‌ها بازی سیستمی ندارند: بعضی از بیمارستان‌ها واقعا فرآیند را مورد اصلاح قرار داده‌اند تا هم نتایج بیماران را بهبود ببخشند و هم از هزینه‌های مدی‌کر با کاهش بازپذیرش بکاهند. اما دیگر بیمارستان‌ها به‌سادگی توانایی خود را بهبود دادند اما با دستکاری برچسب‌هایی که بیماران برای قضاوت عملکرد با آنها دسته‌بندی می‌شدند. [۳۳]

عواقب منفی دیگری هم وجود داشت. از سال ۲۰۱۵ تقریباً سه چهارم بیمارستان‌های گزارش‌شده توسط مدی‌کر مجازات شدند. موضوع قابل توجه این است که بیمارستان‌های مهم آموزشی - که تمایل بیشتری به دیدن بیماران با وضعیت حاد دارند - به طور نامناسبی تحت تاثیر قرار گرفتند. [۳۴] بیمارستان‌هایی نیز در مناطق فقیرنشین بودند؛ جایی که احتمال اینکه بیماران پس از ترخیص اولیه از آنها به‌خوبی مراقبت شود (یا به طور کلی از آنها مراقبت شود)، بسیار کم بود. [۳۵] رسیدن به هدف کاهش پذیرش نه تنها بستگی به گام‌هایی دارد که بیمارستان برای آموزش بیمار و تهیه داروهای لازم برمی‌دارد بلکه وابسته به بسیاری از عوامل دیگر است که بیمارستان بر روی آنها کنترلی ندارد: سلامت جسمی و روحی بیمار، سیستم پشتیبانی اجتماعی و رفتاری. چنین عواملی به مساله پرتکرار دیگری در معیارهای پزشکی اشاره دارد: بیمارستان‌ها به بیماران مختلفی خدمات ارائه می‌دهند، برخی از آنها بیشتر مستعد ابتلا به بیماری هستند و توانایی مراقبت از خود را پس از ترخیص ندارند. برنامه‌های پرداخت مبتنی بر عملکرد تلاش می‌کنند این مورد را با «تعدیل خطر» جبران کنند. اما محاسبات میزان خطر نیز مانند سایر معیارها در معرض سوءاستفاده و دستکاری قرار دارند. در نهایت احتمال اینکه بیمارستان‌های خدمات‌دهنده به بیماران با موقعیت حادث‌تر مجازات شوند، بیشتر است. [۳۶] مانند مدارس که برای عملکرد ضعیف دانش‌آموزان خود در آزمون‌های استاندارد مجازات می‌شدند، با مجازات بیمارستان‌هایی که کمتر از بقیه موفق بوده‌اند نیز معیارهای عملکرد

ممکن است منجر به تشدید نابرابری در توزیع منابع شود- و به سختی می‌توان سهم آنها را در سلامت عمومی مثبت ارزیابی کرد.

### صفحه تعادل

در حال حاضر بیشتر سازمان‌های ارائه‌دهنده خدمات درمانی از معیارها برای بهبود کیفیت اهداف خود استفاده می‌کنند؛ از بهبود نتایج فرآیندهای ویژه تا بهینه‌سازی عملیات کل یک موسسه. این استفاده داخلی از معیارهای عملکرد دارای اهمیت بسیاری در کمک به بیمارستان‌ها و دیگر موسسه‌های پزشکی برای افزایش ایمنی و کارآیی مراقبت‌های پزشکی دارد. اما معیارها در درمان‌ها و نتایجی کاملاً موفق هستند که به طور کامل در نظام پزشکی سازمان کنترل شوند؛ مانند مورد فهرست فرآیندها برای به حداقل رساندن عفونت در خطوط مرکزی. وقتی که نتایج به عوامل بسیار گسترده‌تری وابسته باشند (مانند رفتار بیمار بیرون از مطب دکتر و بیمارستان) تخمین تلاش‌ها و شکست‌های نظام پزشکی دشوارتر می‌شود. موفقیت جیسینجر در مدیریت سلامتی جمعیت با دادن امید بود. این کار اما در شرایطی انجام می‌شود که معیارهای تشخیصی در یک فرهنگ سازمانی بزرگ‌تر نقش ایفا کنند که در آن چنین معیارهایی با رعایت اخلاق حرفه‌ای توسط خود پزشکان تدوین و ارزیابی شود.

استفاده از معیارها برای پاداش عملکرد، چه پولی و چه اعتباری بسیار مشکل‌ساز است. توسل به معیارها در راستای مشوق‌های پولی و رتبه‌های عمومی افزایش یافته است. اینکه آیا به این موارد هزینه‌ها و مزایای مراقبت‌های بهداشتی را اضافه یا کم می‌کنند همچنان یک پرسش بی‌پاسخ است.

## فصل ۱۰

## پلیس

مانند پزشکی، پلیس نیز در دهه‌های اخیر با استفاده از معیارها تغییر کرده است. اینجا نیز ریسک‌پذیری بالاست: سرنوشت شهرها ارتباطی با درک عموم از امنیت ندارد و شهرداران غالباً برای انتخاب مجدد روی توانایی‌های خود در کنترل جرائم یا پایین آمدن میزان آن تمرکز می‌کنند. هنگامی که مردم و سیاستمداران‌شان به امنیت عمومی فکر می‌کنند پلیس را در نظر می‌گیرند که مسئولیت میزان جرائم با اوست. مانند سلامت و ارتباط آن با سیستم پزشکی یا آموزش و ارتباط آن با سیستم مدرسه، امنیت عمومی نیز تا حدی به پلیس و اثربخشی آن وابسته است. این مورد تا اندازه‌ای به سایر عوامل سیستم قضایی نیز بستگی دارد: دادستان‌های عمومی، دادگستری و سیستم‌های کیفری و عفو مشروط. این امر بستگی به گرایش مردم محلی به انجام فعالیت‌های جنایی دارد که این مورد نیز به نوبه خود به عوامل اقتصادی، قومی و فرهنگی وابسته است. [۱] امنیت عمومی نیز ارتباط مستقیمی با سهولت ارتکاب جرم دارد. کاهش برخی جرائم در دهه‌های

اخیر محصول فعالیت‌ها و اقدامات مالکان خصوصی است. با اقدامات دفاعی انجام‌شده توسط این افراد امکان سرقت خودرو، دزدی و دیگر جرائم به طور چشمگیری کاهش یافته است، استفاده از دزدگیرهای اتومبیل و خانه نیز موجب شده ارتکاب این جرائم دشوارتر شود.

جرائم خشونت‌آمیز در ایالات متحده آمریکا در اوایل دهه ۱۹۹۰ به شدت کاهش یافت. درست یا غلط، تغییر روال پلیس سهم بزرگی در این کاهش داشت. این تغییر شامل استفاده از معیارها و بالاتر از آن به شکل کامپاستات<sup>۱</sup> بود. در اینجا موردی وجود دارد که معیارهای تشخیصی برای استفاده داخلی بسیار مفید بوده‌اند اما پس از آن مجدداً استفاده از معیارهای منتشرشده به صورت عمومی در راستای افزایش اعتبار سیاست‌مداران و روسای پلیس موجب بروز انگیزه‌هایی برای بازی و دستکاری اعداد و ارقام و انحراف تلاش‌ها شد. کامپاستات (که در اصل مخفف «آمار کامپیوتری» است) یک سیستم تجزیه و تحلیل و پاسخگویی جرائم است که اولین بار توسط اداره پلیس نیویورک در سال ۱۹۹۴ ایجاد شد. این سیستم که نخستین بار توسط کمیسر ویلیام جی. براتون<sup>۲</sup> معرفی شد، از سیستم‌های اطلاعات جغرافیایی (جی‌آی‌اس) برای ردیابی جرائم استفاده می‌کرد. سیستم کامپاستات جمع‌آوری، تجزیه، تحلیل و نقشه‌برداری از داده‌های جرائم در یک بازه زمانی سریع برای کشف الگوهای جرائم و همچنین برگزاری جلسات هفتگی است که مدیران در این جلسات در مورد نتایج حوزه‌های خود پاسخگو هستند. از داده‌ها برای مشخص کردن نقاطی که جرائم در آنها متمرکزند استفاده می‌شود و نیروهای پلیس بر اساس آن مستقر می‌شوند. در دهه‌هایی که این سیستم در نیویورک معرفی شد تغییرات بسیاری در آن به وجود آمد و در بیشتر شهرهای بزرگ آمریکا مورد استفاده قرار گرفت. [۲] به نظر می‌رسد کامپاستات در کاهش جرائم گزارش‌شده و در حقیقت در

1. Compstat: Computer Statistics

رویکردی پویا در اداره پلیس آمریکا برای کاهش جرم و جنایت در این کشور که ترکیبی از مدیریت، فلسفه و ابزارهای مدیریت سازمانی است (منبع: ویکیپدیا). م

2. William J. Bratton

کاهش جرائم نقش داشته است.

هنوز در شهرها یکی پس از دیگری درباره صحت و اعتبار آمار جرائم پرسش‌هایی مطرح است. از آنجا که کامپاستات یک سیستم اطلاعاتی و نمادین از معیارهاست، به نظر می‌رسد که واقعاً سودمند باشد. اما هنگامی که شهردار، افسر ارشد پلیس را تحت فشار قرار می‌دهد تا پیشرفت را با اعداد و ارقام نشان دهد و این فشار به نوبه خود به رئیس پلیس مناطق وارد می‌شود، آنها به این باور می‌رسند که پیشرفت شغلی آنها وابسته به کاهش مدام جرائم است؛ این پیام به گوش افسران پلیس سطوح پایین‌تر نیز می‌رسد که آنها برای افزایش جرائم گزارش‌شده مجازات خواهند شد. این فشارها می‌تواند به دستکاری و ساخت ارقام بینجامد.

البته، چنین مشکلاتی پیش از ظهور کامپاستات نیز وجود داشت. در سال ۱۹۷۶، روان‌شناس اجتماعی دونالد تی. کمبل (قانون کمبل در فصل یک) اظهار کرد: «اعلام سخت‌گیری رئیس‌جمهور ریچارد نیکسون بر جرائم مهم‌ترین تاثیر را روی شاخص‌های فساد و جنایت داشت که این امر از طریق دست‌کم گرفتن جرم و جنایت و دسته‌بندی آنها به جرائم کمتر جدی حاصل شد»<sup>[۳]</sup> و این کار ادامه یافت. بیشترین میزان گزارش از معیار جرائم مربوط به نظام هماهنگ گزارش جرائم<sup>۱</sup> اداره اف‌بی‌آی<sup>۲</sup> تحقیقات فدرال است. طبق گزارش‌های مربوط به هر شهر، اف‌بی‌آی داده‌های مربوط به چهار جنایت بزرگ خشونت‌آمیز (قتل، تجاوز، حمله، سرقت) و چهار جرم مهم مربوط به اموال (سرقت، دزدی، سرقت وسایل نقلیه و آتش‌زدن) را جمع‌آوری می‌کند. جرائم کوچکتر در شاخص قرار نمی‌گیرند و این شاخص‌ها به طور گسترده منتشر و به عنوان برگه گزارش جرائم در نظر گرفته می‌شوند. هنگامی که جرائم کاهش یابد، مقامات منتخب موفقیت خود را با صدای بلند اعلام می‌کنند. هنگامی که شاخص جرائم بالا می‌رود، سیاستمداران از رقبای خود انتقاد می‌کنند. سیاستمداران بر روسای پلیس فشار می‌آورند تا آمار

1. Uniform Crime Report

2. FBI: Federal Bureau of Investigation

جرائم کاهش یابد و آنها نیز به نوبه خود این فشار را به زیردستان در سلسله مراتب پلیس منتقل می‌کنند.

همه اینها اشتیاق فوق‌العاده‌ای برای نشان دادن پیشرفت در کاهش جرائم با دستکاری اعداد و ارقام ایجاد می‌کند. همان‌گونه که یکی از کارآگاهان شیکاگو توضیح داده است:

«بسیار ساده است.» نخست، افسر پاسخگو می‌تواند عمداً مورد را اشتباه دسته‌بندی کند یا داستان را به‌گونه‌ای تغییر دهد تا هزینه کمتری ثبت شود. به این ترتیب، ورود به خانه می‌شود «تخلف»، ورود به گاراژ می‌شود «خسارت کیفری به اموال» و دزدی می‌شود «اموال گمشده». [۴]

در هر یک از این موارد، آنچه که یک جرم بزرگ بوده بدل به یک جرم کوچک شده و در گزارش نظام هماهنگ جرائم اِیبی‌آی منتشر نشده است. تمایل به دست‌کم گرفتن جرم آنقدر خوب است که اداره پلیس نیویورک منابع قابل توجهی برای ممیزی گزارش‌های دریافتی اختصاص داده و مشخص شده است که ماموران گزارش اشتباه داده و در نهایت مجازات شده‌اند. [۵] اما همه نیروهای پلیس چنین منابعی ندارند یا خواهان این نیستند که با چنین مواردی برخورد کنند.

این مشکل محدود به ایالات متحده آمریکا نیست. در لندن، دفتر شهردار برای پلیس هدف عملکردی کاهش ۲۰ درصدی جرایم را تعیین کرده است. این هدف به زنجیره نیروهای پلیس منتقل شد؛ از افسران عالی‌رتبه تا افسران وظیفه که برای همه آنها شانس ارتقا با رسیدن به هدف ۲۰ درصد ارتباط مستقیم داشت. در سال ۲۰۱۳ یک نیروی افشاگر پلیس گفت که کمیته پارلمانی، دستکاری آمارها را به «بخش ثابت فرهنگ پلیس» بدل کرده است: جرایم بزرگ مانند سرقت به «قاپ‌زنی»<sup>۱</sup> تقلیل می‌یابند و تجاوزها نیز در راستای رسیدن به اهداف تعیین‌شده کمتر از میزان واقعی گزارش می‌شوند. همان‌طور که یک کارآگاه عالی‌رتبه اعلام کرد: «هنگامی که دفاتری مانند دفتر شهردار برای اداره پلیس هدف

تعیین می‌کنند، آنچه که فکر و درخواست می‌کنند ۲۰ درصد کمتر از قربانیان است.» این بدان معناست که جرائم ۲۰ درصد کمتر «توسط افسران ثبت می‌شود.» او افزود: «چنین گزارش‌هایی و تقلیل جرائم یک حقیقت معمول در هر سطح از نیروهای پلیس در انگلستان و ولز است.» سایر کارشناسان تکنیک‌های مختلفی را برای بهبود معیارهای عملکرد توضیح دادند: باور نکردن شکایت‌کنندگان، ثبت چندین حادثه در یک منطقه به عنوان یک جرم و تقلیل وقایع به جرائم کم‌اهمیت‌تر. [۶]

اشتیاق دیگری نیز هست که شاید مشکل‌سازتر باشد؛ این مورد یک معیار کلیدی مهم در موفقیت پلیس است: آمار بازداشت‌ها در حکم معیار اثربخشی. اد پرنز<sup>۱</sup> یکی از کارآگاهان سابق بالتیمور که در بخش‌های قتل و مواد مخدر کار می‌کرد - و بیشتر به دلیل اینکه یکی از تهیه‌کنندگان سریال شنود است شناخته می‌شود - فرآیند «شوخی آماری» را شرح داده است که بدین وسیله مقامات پلیس می‌توانند فعالیت‌های این اداره را به سوی نتایج چشمگیر سوق دهند. پرنز به عنوان یک کارآگاه در بخش مواد مخدر به دنبال ایجاد پرونده‌ای علیه اربابان ارشد مواد مخدر بود اما مافوق او تمایلی به این کار نداشت، زیرا باید نیروی انسانی بسیاری سال‌ها صرف یک بازداشت می‌شد. آنها علاقه‌مند به پیشبرد معیارها بودند و از آنجا که دستگیری پنج نوجوانی که در کنار خیابان مواد می‌فروختند نتیجه آماری بهتری از دستگیری یک سردسته باند مواد مخدر پس از تحقیقات چندساله داشت، آنها طرفدار این بودند که ظرف مدت کوتاهی آمار و ارقام بالایی تولید کنند. از این منظر و از منظر سیاست‌مدارانی که به آنها گزارش می‌دادند، هر بازداشت ارزش یکسانی داشت. بنابراین دوره عملیاتی که بهترین شاخص‌های عملکرد را تولید کرد، تاثیر بسیار اندکی بر فروش مواد مخدر گذاشت. [۷] هنگامی که واحدهای وزنی برابر است، تمایل این است که به دنبال ساده‌ترین موارد برویم. [۸]

در بریتانیا، این فرآیند هدایت منابع پلیس به سوی جرائم ساده‌تر در راستای افزایش و تقویت آشکارسازی به منزله «کم‌کاری»<sup>۲</sup> شناخته

1. Ed Burns

2. Skewing

می‌شود.

بنابراین معیارها نقش سودمندی در اداره پلیس ایفا می‌کنند. اما تلاش برای استفاده از معیارها به مثابه بنیانی برای پاداش و مجازات می‌تواند منجر به استفاده از معیارهایی شود که کمتر قابل اعتماد و حتی آسیب‌زا (غیرمولد) هستند.

## فصل ۱۱

# ارتش

ارتش آمریکا شاید بزرگ‌ترین و پیچیده‌ترین سازمان جهان باشد. دست‌کم از زمان ویتنام و اخیراً در عراق و افغانستان این سازمان تلاش کرده تا از معیارها در مبارزات ضد شورش<sup>۱</sup> استفاده کند. اگر چه بخش کوچکی از ارتش ایالات متحده از معیارها استفاده می‌کند اما مبارزات ضد شورش یک مورد آموزنده با نتایج بزرگ‌تر برای موضوع ماست. ارتش نه تنها از معیارها به نفع مسئولیت‌پذیری و شفافیت، استفاده گسترده‌ای کرده بلکه تلاش‌های پژوهشگران مشغول به کار و پژوهش در دانشگاه‌های نظامی آمریکا و شرکت رند<sup>۲</sup> و وزارت دفاع نیز به دقت مورد بررسی قرار گرفته است. برخی از این پژوهشگران هم سرباز هستند و هم دانشجو و برخی دیگر سوابق تحصیلی معمولی تری دارند. آنچه که مشخصه کار آنهاست ارتباط نزدیک با تجربه واقعی است؛ چه به صورت مشارکت مستقیم در عملیات ضد شورش و چه دسترسی به افسران مشارکت‌کننده. نوشتن یک متن خوب برای سیاست‌گذاران و افسرانی

1. Counterinsurgency (COIN)

2. Rand Corporation

که در آینده به کار گرفته خواهند شد دارای هزینه‌های بسیاری است. در نتیجه، شاید برخی در استفاده و سوءاستفاده از معیارها دارای صداقت و دقت باشند. [۱]

همان‌گونه که تجربه آمریکا در ویتنام نشان می‌دهد، معیارها ممکن است گمراه‌کننده و پیگیری آنها می‌تواند پیامدهای منفی پنهانی داشته باشد. اول اینکه جمع‌آوری اطلاعات چه‌بسا بسیار پرهزینه باشد: سربازان آمریکایی جان‌شان را در راه جست‌وجو برای پیدا کردن اجساد از دست دادند زیرا معیار وزیر دفاع، مک‌نامارا، شمارش اجساد دشمن بود (فصل سه را ببینید). این آمار غالباً در راستای افزایش شانس ارتقای افسران فرمانده بزرگ‌نمایی می‌شود. جریان اطلاعات به ظاهر عینی اما در واقع اشتباه منجر به این می‌شود که سیاست‌گذاران و سیاست‌مداران به اشتباه بهبود در اندازه‌گیری عملکرد را یک پیشرفت واقعی بدانند. [۲]

دیوید کیلکولن<sup>۱</sup> یک سرباز-پژوهش‌گر بود که پیش از رفتن به ایالات متحده آمریکا افسر ارتش استرالیا بود و چند پست کلیدی داشت؛ او استراتژیست عملیات ضد شورش در ارتش آمریکا و وزارت امور خارجه بود و مدتی را در افغانستان و عراق سپری کرده بود. کتابش تحت عنوان «ضد شورش» در حقیقت مقاله‌ای روشنگرانه درباره «اندازه‌گیری پیشرفت در افغانستان» است. «عملیات ضد شورش»، آن‌طور که او به‌سادگی می‌گوید، عبارت است از: «هر آنچه دولت‌ها برای شکست شورش‌ها انجام می‌دهند». [۳] محیطی که با ضد شورش‌ها روبه‌رو می‌شود پیچیده و پویاست: «شورشیان و تروریست‌ها در پاسخ به اقدامات متقابل به سرعت تکامل می‌یابند به نحوی که ممکن است آنچه که یک بار نتیجه‌بخش بوده بار دیگر این‌طور نباشد و دیدگاه‌هایی معتبر برای یک منطقه یا دوره شاید در جای دیگر کارایی نداشته باشد». از این رو، کیلکولن تأکید می‌کند که معیارها باید با ویژگی‌های یک مورد سازگار باشند: معیارهای استاندارد که از

1. David Kilcullen

جنگ‌های گذشته در مکان‌های دیگر به دست آمده به‌سادگی عمل نمی‌کنند، هرچند استفاده از بهترین معیارهای عملکرد بر اساس تجربه نیز نیازمند قضاوت و داوری است:

تفسیر شاخص‌ها بسیار مهم است و نیاز به قضاوت کارشناسان خبره دارد. صرفاً شمارش وقایع یا تجزیه و تحلیل کمی و آماری کافی نیست، تفسیر یک فعالیت کیفی است که بر مبنای آشنایی با محیط انجام می‌شود و باید توسط کارکنان با تجربه‌ای انجام شود که مدت‌های طولانی در آن محیط کار کرده و به اندازه کافی از روندهایی پیشین برای مقایسه با رویه‌های جدید آگاه باشند. مثلاً ممکن است این روند برای افرادی مشغول در ماموریت‌های کوتاه‌مدت کشور مشخص نباشد. [۴]

کیلکولن توضیح می‌دهد که چرا بسیاری از معیارهای استاندارد می‌تواند فریب‌دهنده باشند و باید از آنها دوری کرد، از جمله شمارش اجساد یا شمارش «فعالیت‌های قابل توجه»<sup>۱</sup> که به معنای حوادث خشونت‌آمیز علیه نیروهای ضد شورش است. فرض معمول این است که هرچه تعداد چنین برخوردهای خشونت‌آمیزی کمتر باشد، بهتر است. کیلکولن توضیح می‌دهد که لزوماً این‌گونه نیست زیرا «خشونت در مناطق جنگی زیاد و در مناطق تحت کنترل دولت کم است. هرچند در مناطق تحت کنترل دشمن نیز میزان خشونت زیاد نیست، بنابراین سطح پایین خشونت به ما می‌گوید که منطقه تحت کنترل افرادی است، اما نمی‌گوید چه افرادی.» او همچنین درباره استفاده از «معیارهای ورودی»<sup>۲</sup> هشدار می‌دهد زیرا این معیارها آنچه را که ارزش و متحدان‌شان انجام می‌دهند، محاسبه می‌کند و ممکن است این اقدامات کاملاً متمایز از نتایج باشد:

«معیارهای ورودی، شاخص‌هایی است بر اساس سطح تلاش‌های

1. SIGACTs: significant activity

اصطلاحی که در جنگ آمریکا با عراق شامل کلیه حملات گزارش‌شده علیه نیروهای ائتلاف، نیروهای امنیت عراق و افراد غیرنظامی می‌شد. م

2. Input metrics

ما و می‌تواند متمایز از تاثیر تلاش‌های ما باشد. مثلا معیارهای ورودی شامل تعداد کشته‌های دشمن، تعداد نیروهای آموزش‌دیده، دوستان، تعداد مدارس و کلینیک‌های ساخته‌شده، میزان جاده‌های کامل‌شده و ... است. این شاخص‌ها به ما می‌گویند ما چه کاری انجام می‌دهیم اما تاثیری را که بر آنها داریم، نمی‌گویند. برای فهم این تاثیر، باید به معیارهای خروجی یا بهتر است بگوییم به معیارهای نتایج توجه کنیم (چگونه نیروهای دوست سه ماه پس از آموزش خدمت می‌کنند، یا مثلا چه تعداد مدرسه و کلینیک پس از یک سال مورد استفاده قرار می‌گیرند). معیارهای نتیجه اثر واقعی و درک‌شده فعالیت‌ها و اقدامات ما را بر امنیت، ایمنی و رفاه جمعیت ردیابی می‌کنند.» [۵]

یافتن معیارهای جدید اغلب نیازمند شناخت عمیق از شرایط محلی است. مثلا قیمت بازار سبزیجات خارجی (غیرمحلی) را در نظر بگیرید که از نظر تعداد اندکی از افراد خارجی در حکم شاخصی از صلح و رفاه مردم در نظر گرفته می‌شود. با این حال، کیلکولن توضیح می‌دهد که چرا این شاخص‌ها می‌تواند مفید باشد:

افغانستان یک اقتصاد کشاورزی دارد و دارای تنوع زراعی قابل توجهی است. با توجه به اقتصاد بازار آزاد تولید محصولات کشاورزی در افغانستان، به نظر می‌رسد عوامل خطر و هزینه (هزینه-فرصت رشد یک محصول، خطر حمل و نقل آنها در جاده‌های ناامن، خطر فروش آنها در بازار و حمل پول به خانه) به طور خودکار به قیمت میوه و سبزیجات وارد می‌شود. بنابراین نوسانات کلی قیمت بازار شاید یک معیار جایگزین برای اعتماد به نفس عمومی و امنیت درک شده باشد. به طور خاص، سبزیجات خارجی-آنهايي که در خارج از یک منطقه خاص رشد می‌کنند و خطر بیشتری در حمل برای فروخته شدن در آن محل را دارند- می‌تواند یک نشان‌گر سودمند باشد. [۶]

بنابراین، توسعه معیارهای معتبر موفقیت و شکست نیازمند داشتن سطح خوبی از دانش محلی است؛ دانشی که ممکن است در سایر موقعیت‌ها و مکان‌ها بی‌فایده باشد و موجب آزرده‌گی کسانی شود که به دنبال الگوها و فرمول‌های جهانی هستند. بخش دشوار، دانستن آنچه که شمارش می‌شود و اعداد شمارش‌شده معنادار در یک حوزه است. درس‌های بزرگ‌تری نیز از ارزیابی ضد شورش توسط بن کانابل<sup>۱</sup> تحلیلگر شرکت رند در کتابی با عنوان «آغوش مه‌آلود جنگ: ارزیابی و معیارها در ضد شورش<sup>۲</sup>» ارائه شده است. او می‌نویسد: «تهیه یک مدل عملی و متمرکز برای ارزیابی ضد شورش دشوار است (اگر غیرممکن نباشد) زیرا محیط‌های پیچیده ضد شورش را نمی‌توان با یک فرآیند متمرکز که داده‌ها را از زمینه محلی آنها دور می‌کند، به روشنی تفسیر کرد.» بنابراین، «اطلاعات می‌تواند از مکانی به مکان دیگر و از زمانی به زمان دیگر معانی متفاوتی داشته باشد.» مشکل ناشی از «ناسازگاری میان عملیات پیچیده و غیرمتمرکز ضد شورش و ارزیابی متمرکز و خارج از متن است.» [۷]

این دغدغه‌ها فراتر از قلمروی نظامی است: ما تلاش می‌کنیم که معیارهای عملکرد را برای هر محیط یا سازمان پیچیده‌ای ایجاد کنیم که یا منحصربه‌فرد است یا با محیط‌ها و سازمان‌های دیگر تفاوت دارد در حالی که اندازه‌های استاندارد شده عملکرد، نادرست و گمراه‌کننده خواهند بود. با این حال، گرایش به ایجاد معیارهای عملکردی که «شفاف» و «پاسخگو» باشند وجود دارد و معمولاً به استفاده از معیارهای استاندارد و متمرکز تبدیل می‌شود زیرا این معیارهای ساده‌تر توسط مقامات مافوق و افراد دور از عملیات درک می‌شود. علاوه بر این، یکی دیگر از مطالعات اخیر رند یادآور می‌شود که مشاهداتی که از طریق اندازه‌های کمی به آنها ابلاغ می‌شود «تجربی» قلمداد می‌شود، در حالی که مشاهدات منتقل‌شده به صورت کیفی کمتر قابل اطمینان تلقی می‌شود، با وجود این واقعیت که «در عمل، بسیاری از معیارهای

1. Ben Connable

2. Embracing the Fog of War: Assessment and Metrics in Counterinsurgency

کمی مورد استفاده در ارزیابی‌ها منعکس‌کننده مشاهدات تعصب‌آمیز از گزارش‌های ارائه‌شده است.» [۸]

کانابل عملیات ضد شورش را «هم علم و هم هنر، اما بیشتر هنر» توصیف کرده [۹] که مربوط به مدیریت بسیاری از موقعیت‌های پیچیده دیگر است. گرایش به این سمت است که با عملیات ضد شورش در حکم یک علم ناب و قابل اندازه‌گیری رفتار کنیم که ضرورتاً بخش بزرگی از آن هنر است و نیاز به قضاوت مبتنی بر تجربه دارد.

## فصل ۱۲

## کسب‌وکار و مالیه

چه زمانی پرداخت مبتنی بر عملکرد مناسب است و چه زمانی نیست؟ «اما مطمئناً» شاید شما فکر کنید «اگر جایی وجود داشته باشد که پرداخت مبتنی بر عملکرد مناسب باشد، آن حوزه کسب‌وکار و تجارت است.» به هر حال، کسب‌وکارها برای ایجاد درآمد است و مردم برای کسب درآمد در آنها کار می‌کنند. به نظر می‌رسد مدیران کسب‌وکار سعی دارند تا بزرگ‌ترین تلاش‌های کارمندان خود را با پاداش دادن به سهم‌شان در سودآوری شرکت پیوند بزنند.

شرایطی وجود دارد که پرداخت مبتنی بر عملکرد این وعده را تحقق می‌بخشد: هنگامی که کاری که باید انجام شود تکراری و غیرخلاقانه و شامل تولید و فروش کالاها و خدمات استاندارد باشد، هنگامی که امکان انتخاب بسیار کمی بر آنچه که انجام می‌شود وجود داشته باشد، زمانی که رضایت ذاتی اندک باشد، هنگامی که عملکرد منحصر به تلاش فردی باشد تا تلاش تیمی و زمانی که کمک، تشویق و راهنمایی دیگران بخش مهمی از یک کار نباشد.

برای نیروهای فروش [۱]، مشاغل معمولی، فردی و بسیار متمرکز که

دارای خروجی‌های استاندارد بدون مسئولیت‌های گسترده است، امکان دارد پاداش عملکرد اندازه‌گیری شده به‌خوبی پرداخت شود. به‌طور خلاصه، همان‌طور که یک جامعه‌شناس گفته است: «پاداش‌های بیرونی یک عامل تعیین‌کننده بسیار مهم در رضایت شغلی در میان کارکنانی است که پاداش‌های داخلی (ذاتی) برای‌شان تقریباً غیر قابل دسترس است.» [۲] اینها نوعی از وظایفند که تیلوریسم (فصل سه را ببینید) برای آنها طراحی شده است. تعداد این مشاغل در هر جامعه از جمله جوامع پیشرفته از نظر تکنولوژی زیاد است اما در زمان ما با پیشرفت فناوری‌های رباتیک و هوش مصنوعی چنین مشاغلی به‌سرعت در حال کم شدن است. [۳]

اما واقعیت مهم این است که بیشتر مشاغل بخش خصوصی با این معیارها مطابقت ندارد. از این رو، پرداخت مبتنی بر عملکرد به‌طور مستقیم نه تنها برای این مشاغل نامناسب است که ممکن است نامولد (آسیب‌زا) نیز باشد. افراد می‌خواهند به‌خاطر عملکردشان از آنها قدردانی شود؛ هم از نظر به رسمیت شناخته شدن عملکردشان و هم از نظر پاداش. اما میان پیشرفت بر اساس طیف وسیعی از کیفیت‌ها و پاداش مستقیم بر اساس کمیت‌های اندازه‌گیری شده خروجی‌ها تفاوت وجود دارد. برای بسیاری از کارکنان، مشارکت در شرکت‌هایی که کار می‌کنند، شامل بسیاری از فعالیت‌های واقعی اما نامشخص و مبهم است؛ ارائه ایده‌ها و روش‌های جدید انجام کارها، تبادل ایده‌ها و منابع با همکاران، مشارکت در کار تیمی، راهنمایی و هدایت زیردستان، ارتباط با تامین‌کنندگان و مشتریان و موارد دیگر. درست این است که به‌چنین فعالیت‌هایی از طریق ارتقا و جوایز پاداش داده شود، حتی اگر مستندسازی آن دشوارتر باشد و نیازمند قضاوت و داوری بیشتر توسط افرادی باشد که درباره پاداش تصمیم می‌گیرند. مشکل در تعیین و تخصیص اعداد به عملکرد و یا رتبه‌بندی افراد بر مبنای یک مقیاس نیست. مشکل از آنجا شروع می‌شود که مقیاس بیش از اندازه تک‌بعدی باشد و تنها خروجی‌های اندکی که به‌سادگی قابل اندازه‌گیری است به‌عنوان

یک مقیاس استاندارد معرفی شود. در حقیقت، شواهد دانشگاهی پرداخت مبتنی بر عملکرد برای سنجش عملکرد مدیران عامل و دیگر پرسنل به اندازه کافی نگران‌کننده است تا آنجا که برخی پژوهشگران رفتار سازمانی، پیشنهاد حذف آن را داده‌اند. برخی شرکت‌ها نیز بر این اساس عمل می‌کنند. دن کابل<sup>۱</sup> و فریک ورمیولن<sup>۲</sup> از مدرسه کسب‌وکار لندن یادآور بسیاری از مشکلاتی می‌شوند که ما آنها را یافتیم: تاثیر بحران پرداخت مبتنی بر عملکرد بر خلاقیت، تمایل به تحریف کتاب‌ها، نقص اجتناب‌ناپذیر ابزارهای اندازه‌گیری، دشواری در تعریف عملکرد بلندمدت و تمایل به انگیزه‌های بیرونی برای کنار زدن انگیزه‌های درونی و ذاتی. آنها به این نتیجه رسیدند که لغو کارمزد برای مدیران برتر و جایگزین کردن آن با حقوق ثابت می‌تواند مفیدتر باشد. آنها حتی پا را فراتر می‌گذارند و می‌گویند شاید مدیران رده‌بالا با انگیزه‌ای بیرونی به شما انگیزه ندهند: با این حال، هرچه میزان پرداخت بیشتر تغییر کند و با عملکرد اندازه‌گیری شده بیشتر در ارتباط باشد احتمال بیشتری نیز وجود دارد که این افراد در رأس شرکت شما قرار بگیرند [۴] و دست‌کم یکی از شناخته‌شده‌ترین سرمایه‌گذاران انگلیس به نام نیل وودفورد<sup>۳</sup> که شرکت مدیریت سرمایه‌گذاری وودفورد<sup>۴</sup> را به ارزش ۱۴/۳ میلیارد پوند مدیریت می‌کند برای مدیران شرکتش جوایز و پاداش را حذف کرده و به جای آن پرداخت ثابت را افزایش داده است با این استدلال که ارتباط کمی میان پاداش و عملکرد وجود دارد. [۵]

رتبه‌بندی اجباری که طی آن به مدیران آموزش داده می‌شود کارمندان خود را با یکدیگر مقایسه و ارزیابی کنند، یکی دیگر از نشانه‌های وجود ثبات معیارهاست. به نظر می‌رسد این مورد «سخت» و «عینی» باشد، اما در نهایت آسیب‌زا و نامولد است. بررسی بیش از ۲۰۰ نیروی انسانی متخصص در شرکت‌های بزرگ در سال ۲۰۰۶ نشان می‌دهد که «حتی

1. Dan Cable

2. Freek Vermeulen

3. Neil Woodford

4. Woodford Investment Management

اگر بیش از نیمی از شرکت‌ها از رتبه‌بندی اجباری استفاده کنند باز هم این رویکرد به تولید کمتر، نابرابری، تردید، کاهش اشتغال کارمندان، کاهش همکاری، آسیب به روحیه و بی‌اعتمادی به رهبری (مدیریت) می‌انجامد.» [۶]

تعداد فزاینده‌ای از شرکت‌های فناوری از تاثیر مخرب رتبه‌بندی‌های عملکرد بر اکثر کارکنان خود آگاه بوده و پاداش عملکرد را حذف کرده‌اند. آنها با جایگزین کردن حقوق پایه بالاتر همراه با سهام یا جایگزین‌های سهام به کارمندان‌شان گرایش و انگیزه ملموسی به شکوفایی شرکت در بلندمدت می‌دهند (در حالی که پاداش‌های ویژه‌ای نیز به افرادی که عملکردهای درخشانی داشته‌اند نیز می‌دهند [۷]).

با این حال شرکت‌های دیگر، رتبه‌بندی‌های سالانه را در حمایت از «انبوه‌سپاری (جمع‌سپاری)»<sup>۱</sup> کاهش می‌دهند اما همچنان به جمع‌آوری داده‌ها می‌پردازند که این کار توسط سرپرستان، همکاران و مشتریان داخلی انجام می‌شود و دائماً بازخورد آنلاین درباره عملکرد کارمندان ارائه می‌دهد. این کار ممکن است از چاله در آمدن و افتادن در چاه باشد زیرا کارمندان دائماً برای خوشامد و تعریف دیگران در حال ایفای نقش هستند، در حالی که از نظارت همه‌جانبه بر فعالیت‌های خود به ستوه آمده‌اند [۸]؛ یک احتمال ویران‌شهری<sup>۲</sup> که در زمان دایره<sup>۳</sup> نوشته دیو اِگِرز<sup>۴</sup> در سال ۲۰۱۴ ارائه شده است. با این حال، پیشرفت در فناوری اطلاعات، نظارت بر این یا دیگر شاخص‌های عملکرد کارکنان را ساده‌تر کرده است و چه‌بسا پیوند دادن پرداخت مبتنی بر عملکرد و سوسه‌انگیز شده باشد؛ چه به صورت مقاطعه‌کاری و پاداش و چه به صورت کمیسیون [۹]؛ البته با وجود شواهد خطرناکی مبنی

1. Crowdsourced

انبوه‌سپاری یا جمع‌سپاری ترکیبی از دو کلمه جمعیت و برون‌سپاری به معنای برون‌سپاری به انبوه مردم است. جمع‌سپاری در تلاش برای جذب ایده‌های جدید در فرایند نوآوری بسیاری از بنگاه‌ها امروزه در حال گسترش حوزه‌های مدیریت دانش و برون‌سپاری خلاقیت است (منبع: ویکی‌پدیا). م

2. Dystopian

متضاد یوتوپیایی (آرمان‌شهر / مدینه فاضله)؛ جایی که همه تصاویر ارائه‌شده از آن بد و ناخوشایند است. م

3. The Circle

4. Dave Eggers

بر اندازه‌گیری‌های کوتاه‌بینانه، دلسرد کردن کارهای گروهی و از بین رفتن نوآوری.

بخش عمده‌ای از اختلال عملکرد شرکت‌ها ناشی از برنامه‌های پرداخت مبتنی بر عملکرد است که برای اندازه‌گیری یک نتیجه واحد به طور محدود طراحی شده‌اند. از این رو، مشکلات هم در بالا و هم در پایین نردبان شرکت‌ها پیش می‌آید. نمونه چشمگیر آن میان مدیران ارشد کارخانه داروسازی میلان است. اگرچه این کارخانه در میان بزرگ‌ترین شرکت‌های داروسازی آمریکایی نیست (یازدهم در درآمد و شانزدهم در ارزش بازار)، اما دومین رتبه در دادن پاداش است؛ طی پنج سال تا پایان دسامبر ۲۰۱۵ تنها به سه مدیر ارشد خود بالغ بر ۷۰ میلیون دلار پرداخت کرده و در همین مدت، ارزش سهام این کارخانه ۱۵۵ درصد رشد داشته است. در سال ۲۰۱۴ هیات‌مدیره این کارخانه یک برنامه پرداخت پاداش به مدیران ارشد شرکت تدوین کرد که طبق آن اگر سود شرکت هر سال ۱۶ درصد رشد داشته باشد، آنها پاداش بسیار خوبی دریافت خواهند کرد؛ این بسیار فراتر از انتظارات منطقی از شرکتی است که داروهای عمومی تولید می‌کند و عموماً به عنوان یک «بازار بالغ» به حساب می‌آید که وجود رقابت به سودهای متوسط می‌انجامد.

بزرگ‌ترین مرکز سودآوری میلان اپی‌پن<sup>۱</sup> است؛ وسیله‌ای قلم مانند که اپی‌نفرین (آدرنالین) را به‌سادگی به پوست تزریق می‌کند تا با شوک‌های شدید حساسیت‌زا مقابله کند. از آنجا که هر تزریق مدت‌زمان کوتاهی می‌خواهد و کودکان نیز در معرض خطر چنین شوک‌هایی قرار دارند، بچه‌ها -هم در خانه و هم در مدرسه- و همچنین خانواده‌های دارای یک یا چند عضو در معرض خطر نیز نیازمند داشتن چند قلم از آن هستند. از آنجا که پس از گذشت ۱۲ تا ۱۸ ماه اثر دارو از بین می‌رود، قلم‌ها باید دائماً تجدید شوند. این میلان نبود که اپی‌پن را ساخت بلکه شرکت دیگری این قلم را در سال ۱۹۸۷ به بازار عرضه کرد. میلان در سال ۲۰۰۷ حق تولید این قلم را خرید اما از آنجا که هیچ رقیب

1. EpiPen

موثری در بازار نبود انحصار تزریق اپی نفرین را از آن خود کرد. در سال ۲۰۱۱ میلان یکی از مدیران ارشدش، هدر برش<sup>۱</sup> را به سمت قائم مقام ارشد اجرایی تا ژانویه ۲۰۱۲ ارتقا داد. از سال ۲۰۰۹ تا ۲۰۱۳ شرکت میزان قیمت دو بسته قلم را از ۱۰۰ دلار به ۲۶۳ دلار افزایش داد؛ سپس در می ۲۰۱۴ (درست همان زمان که سیستم انگیزشی جدید برای مدیران ارشد آغاز شده بود) قیمت دو برابر شد و به ۴۶۱ دلار رسید، درست پیش از آنکه دوباره در می ۲۰۱۵ به ۶۰۸ دلار برسد. [۱۰] در تابستان ۲۰۱۶، افزایش قیمت میلان برای این وسیله ضروری - که نه فقط توسط بزرگسالان که (به لطف کمپین بازاریابی میلان) توسط بسیاری از کودکان در سن مدرسه نیز استفاده می شود - منجر به خشم و اعتراض عمومی و یک جلسه استماع کنگره شد. چند سناتور از وزارت دادگستری خواستند تا صورت حساب های این شرکت را مورد بررسی قرار دهد.

چگونه سرمایه گذاران میلان از ایجاد انگیزه برای مدیران ارشد خود سود بردند؟ هنگامی که برش به مدیرعاملی شرکت ارتقا یافت، ارزش سهام شرکت ۲۲ دلار بود. در ژوئن ۲۰۱۵ این ارزش به ۷۳ دلار رسید. اما در نتیجه اعتراضات عمومی علیه شرکت و استماع های کنگره و تحقیقات وزارت دادگستری در اکتبر ۲۰۱۶ ارزش هر سهم به ۳۶ دلار سقوط کرد. به این ترتیب تمرکز تک بُعدی مدیران ارشد روی معیارهای سودآوری کلان منجر به از بین رفتن اعتبار شرکت شد.

در همان زمانی که برنامه پرداخت مبتنی بر عملکرد میلان برای مدیران ارشد در حال از بین بردن شرکت داروسازی بود، یک شرکت بزرگ دیگر نیز با نسخه خود در طرح پرداخت مبتنی بر عملکرد در حال سقوط بود. این شرکت نه در بالای نردبان سازمانی که در پایین آن قرار داشت؛ انگیزه های آن مشوق های پولی نبود بلکه فشار آوردن بر افراد برای رسیدن به اهداف عملکرد بود.

این چیز است که اتفاق افتاد. ولز فارگو<sup>۲</sup> یک بانک بزرگ آمریکایی

1. Heather Bresch

2. Wells Fargo

بود که در شرایط سخت اقتصادی کار می‌کرد. هیات‌مدیره بانک مرکزی نرخ بهره را به کمترین میزان ممکن رساند و همین موضوع کار را برای بانک‌های آمریکایی دشوار ساخت زیرا آنها از وام‌هایی که می‌دادند سود می‌کردند. در سال ۲۰۱۱، این شرکت برای افزایش سودآوری، تشویق به «فروش جانبی»<sup>۱</sup> کرد: برای کارکنان خود سهمیه تعیین کرد تا توجه مشتریان علاقه‌مند را به یک یا چند محصول دیگر (مثلاً حساب سپرده) و خدمات اضافی مانند پوشش چک بی‌محل یا کارت‌های اعتباری که برای بانک بسیار سودآور بود، جلب کنند. نرسیدن به سهمیه به معنای ساعات کار اضافی بدون پرداخت و تهدید به از دست دادن شغل بود (شاید الهام‌بخش آنها الکس بالدوین<sup>۲</sup>، شخصیت فیلم گلن‌گاری گلن راس<sup>۳</sup> بود؛ مدیری که نیروهای فروش خود را بر اساس مسابقات فروش خود راهنمایی می‌کرد: جایزه نفر اول، کادیلاک ال‌دورادو؛ جایزه نفر دوم، مجموعه چاقوهای استیل و جایزه نفر سوم این است که اخراج شود. عکس می‌گیرید؟). اما سهمیه‌ها بسیار بیشتر از تعداد محدود مشتریانی بودند که روزانه به بانک می‌آمدند. کارمندان بانک ولز فارگو برای رسیدن به سهمیه‌های تعیین‌شده به کلاهبرداری‌های کوچک متوسل شدند، مثلاً تعدادی پین<sup>۴</sup> برای ایجاد کارت‌های بدهی و حساب کاربری مشتریان بدون اطلاع آنها- به صورت آنلاین ایجاد کردند. این امر هدف مدیریت ولز فارگو نبود: آنها می‌خواستند کارمندان‌شان برای مشتریان و با اطلاع و تمایل آنها حساب‌های قانونی ایجاد کنند. با آشکار شدن شواهدی مبنی بر سوءرفتار و اقدامات غیرقانونی، ولز فارگو در حدود ۵۳۰۰ کارمند خود را اخراج کرد اما موج سوءرفتار یک واکنش قابل پیش‌بینی نسبت به سهمیه‌های عملکرد بود که مدیران شرکت برای کارمندان خود تعیین کرده بودند.

پس از انتشار خبر رسوایی بزرگ در سپتامبر ۲۰۱۶ ولز فارگو توسط دفتر

1. Cross-selling

2. Alex Baldwin

3. Glengarry Glen Ross

4. PIN: Personal Identification Number

حمایت مالی از مصرف‌کننده به ۱۰۰ میلیون دلار، توسط دادستان شهر لس‌آنجلس به ۵۰ میلیون دلار و از سوی اداره کنترل ارز به ۳۵ میلیون دلار جریمه محکوم شد. خسارت به شرکت فقط پولی نبود، بلکه اعتباری هم بود. ارزش سهام ولز فارگو در اواخر آگوست به ۵۰ دلار و در اواخر سپتامبر به ۴۳ دلار سقوط کرد. یک بار دیگر، پاداش و تنبیه برای عملکرد اندازه‌گیری شده نتیجه معکوس داشت. [۱۱]

موارد میلان و ولز فارگو نمونه‌های جدیدی از یک الگوی قدیمی و رایج هستند که سیاست‌های پرداخت مبتنی بر عملکرد آنها منجر به این می‌شود که کارکنان درگیر فعالیت‌هایی شوند که در درازمدت آسیبی جدی به اعتبار شرکت می‌زند. [۱۲]

آیا این مشکل مربوط به طبیعت انسان است و یا ترویج باور پرداخت مبتنی بر عملکرد اندازه‌گیری شده؟ به عبارت دیگر: آیا مفهوم عوامل خودمحور یک واقعیت زندگیست یا با ایدئولوژی مدیریتی مبتنی بر استفاده از پاداش‌های بیرونی بر اساس مدل‌های ساده رفتار انسانی تشدید شده است و پیش‌گویی‌های خود را تحقق می‌بخشد؟ گاهی نحوه برخورد مدیران و کارمندان با شرکت خود بر شیوه تفکر آنها تاثیرگذار است، به نحوی که آنها به روش‌های خودخواهانه و کوتاه‌بینانه که توسط تقلیل‌دهنده‌ترین نسخه‌های نظریه کارفرما-کارگزار ارائه می‌شوند با فریب و حيله عمل می‌کنند. [۱۳] در حقیقت ممکن است شرایطی ایجاد شود که در آن مدیران و کارمندان درباره چگونگی کار شاخص‌های عملکرد هم اطلاعات زیادی داشته باشند و هم در بهترین موقعیتی باشند که بتوانند شاخص‌ها را به نفع خود دستکاری کنند و به احتمال زیاد این کار را انجام دهند. [۱۴] این موارد را در نظر بگیرید: دنیس کوزولوفسکی<sup>۱</sup> مدیرعامل شرکت تیکو<sup>۲</sup>، برنارد ایبرز<sup>۳</sup> مدیرعامل شرکت ووردکام<sup>۴</sup> و جان ریگاس<sup>۵</sup> مدیرعامل شرکت آدلفیا؛ همه این

1. Dennis Kozlowski  
2. Tyco  
3. Bernard Ebbers  
4. WorldCom  
5. John Rigas  
6. Adelphia

افراد اوایل دهه ۲۰۰۰ به زندان رفتند زیرا از اطلاعات و دانش دقیق خود درباره معاملات شرکت‌های خود و در راستای منافع خویش اقدام به دستکاری معیارهای عملکرد کردند و به همین دلیل مجازات شدند. [۱۵] واکنش به این رسوایی‌ها منجر به تصویب قانون ساربنز-اکسلی<sup>۱</sup> در سال ۲۰۰۲ شد که به دنبال تقویت پاسخگویی شرکت‌ها بود که بخشی از این مسئولیت در اختیار هیات‌مدیره شرکت‌های دولتی بود که به طور قانونی مسئول صحت صورت‌حساب‌های مالی هستند. هرچند عمل به این قانون موجب افزایش هزینه‌های شرکت‌ها شده اما ممکن است موجب افزایش اعتماد عمومی به صحت گزارش‌های مالی شود و شواهدی از مزایای شفافیت ارائه کند. به هر حال افزایش پاسخگویی قانونی هر یک از اعضای هیات‌مدیره موجب تحمیل هزینه‌هایی شد که قابل اندازه‌گیری برای اقتصاددانان نیست. یک مشاور هیات‌مدیره شرکت‌های فورچون ۵۰۰<sup>۲</sup> (که به دلایل مشخص نمی‌خواست نامش فاش شود) به من گفت: «پس از تصویب قانون ساربنز-اکسلی اعضای هیات‌مدیره برای اطمینان از صحت و درستی گزارش‌های مالی شرکت آنقدر روی این موضوع تمرکز کرده‌اند که زمان و تمایل اندکی برای انجام وظایف اصلی خود به عنوان هیات‌مدیره دارند؛ تفکر استراتژیک درباره آینده درازمدت شرکت! بنابراین آنچه که اندازه‌گیری و به‌طور بالقوه مجازات می‌شود، انجام خواهد شد.»

## بحران مالی

بحران مالی ۲۰۰۸ دلایل زیادی داشت که برخی از آنها ناشی از تلاش برای جایگزینی معیارهای استاندارد برای داوری بر اساس دانش محلی بود که خود به دلیل تأثیرات پرداخت مبتنی بر عملکرد تشدید شده بود. [۱۶]

از آنجا که شرکت‌ها از جمله شرکت‌های مالی، دارایی‌شان بیشتر و متنوع‌تر می‌شود به لایه‌های جدید مدیریتی برای نظارت و هماهنگی

1. Sarbanes Oxley

2. Fortune 500

واحدهای متفاوت خود نیازمندند. از دیدگاه مدیریت ارشد، تنوع مدیریت به این معناست که مدیران در حال مدیریت ارزیابی دارایی‌هایی هستند که آشنایی کمی با آن دارند. این امر منجر به جست‌وجوی اندازه‌های استاندارد عملکرد در سازمان‌های بزرگ با بخش‌های مختلف شد. دارایی‌های ضمنی این سازمان‌ها عبارتند از: اطلاعاتی که از نظر عددی قابل اندازه‌گیری و تنها نوع دانش مورد نیاز هستند، داده‌های عددی که می‌توانند جایگزین دیگر اشکال تحقیق شوند، دقت عملی (فرض بر فرمول‌های احتمالی و نه پژوهش تجربی) که می‌تواند جایگزین دانش عملی درباره دارایی‌های اساسی باشد. در بحران مالی نقش مدیران مالی که در تحلیل و دستکاری داده‌های معیار مهارت داشتند اما دانش یا تجربه «واقعی» از چیزهایی که ساخته یا موارد قابل معامله نداشتند، رو به افزایش بود. همان‌طور که نایل فرگوسن<sup>۱</sup> گفته است: «آنهایی که می‌خواهند خدایان را نابود کنند، ابتدا به آموزش ریاضیات می‌پردازند.» در اینجا به شکلی خاص آنچه را که منجر به بحران سال ۲۰۰۸ شد، مرور می‌کنیم. به طور سنتی، بانک‌ها (سرمایه‌گذاران خصوصی) به افرادی که با آنها ارتباط مستقیم داشتند، پیشنهاد وام رهن مسکن داده بودند. از این رو آنها در مسند قضاوت درباره افراد بودند؛ اینکه چه کسی ارزش دادن اعتبار دارد و چه کسی ندارد. البته آنها انگیزه برای قضاوت را نیز داشتند: از آنجا که بانک (یا سرمایه‌گذار) به دادن وام رهنی خود ادامه می‌داد، درآمد آنها در آینده نیز به وام رهنی مسکن متکی بود.

تغییر تقریباً در سال ۲۰۰۰ آغاز شد و تا سال ۲۰۰۸ این سیستم با یک سیستم جدید دیگر جایگزین شد. تغییرات در آیین‌نامه‌های سرمایه بانک‌ها موجب شده است تا تهیه و نگهداری وام رهن‌های مسکن سنتی نسبت به اوراق بهادار دیگر که شامل هزاران رهن هستند کمتر سودآور باشد. [۱۷] در حال حاضر وام‌های رهنی مسکن نه از بانک که از یک شرکت کارگزاری وام مسکن نشات گرفته‌اند که از تعداد وام‌های پرداخت‌شده بدون بهره مالی که موجب بقای آنها در بلندمدت می‌شد،

1. Niall Ferguson

کسب درآمد می‌کردند.

بنیان‌گذاران وام رهنی مسکن مانند کانتری‌واید<sup>۱</sup> به افرادی که به دنبال خریدن خانه بودند وام می‌دادند، سپس این وام‌ها را به صورت بسته‌های هزارتایی بسته‌بندی می‌کردند و آنها را به یک بانک مانند برادران لمان<sup>۲</sup> می‌فروختند. از آنجا که بنیان‌گذاران وام رهنی مسکن علاقه‌ای به بقای وام‌هایی که می‌دادند نداشتند، به طور فزاینده‌ای اقدام به دادن وام‌های «با سند کم» یا «بدون سند» کردند؛ بدین معنا که از وام‌گیرندگان خواسته می‌شد تا هیچ مدرکی مبنی بر اینکه اقدام به بازپرداخت وام می‌کنند، ارائه ندهند. اما بانک چنین چیزی را نپذیرفت، بانک یک «اوراق با پشتوانه وام رهنی» ایجاد کرد - نوعی اوراق مشمول نرخ بهره - که با وام‌های رهنی مسکن تضمین و به سرمایه‌گذاران فروخته می‌شد. با مشورت آژانس‌های رتبه‌بندی (مانند مودی<sup>۳</sup>) وام‌های رهنی مسکن با کیفیت خوب را هم از گیرندگانی که احتمالاً وام‌شان را پرداخت می‌کنند و هم از آنهايي که مشکوک به بازپرداخت هستند با یکدیگر ترکیب می‌کنند تا بیشترین سود را از این اوراق با پشتوانه وام رهنی دریافت کنند [۱۸] و «سهام‌ها» را با درجه‌ای از خطر به نرخ‌های مختلف بهره مبدل سازند. پشت تمام اینها باور به وجود تنوع در بخش مالی برای مراقبت دقیق از هر دارایی بود. باور بر این بود که اگر یک مجموعه از دارایی‌ها با هم جمع شود، دیگر لازم نیست درباره دارایی‌ها به صورت جداگانه اطلاعات زیادی داشته باشید یا درباره بقای آنها قضاوت کنید.

ابزارهای مالی پیچیده ریاضیاتی جدید مانند سوپا نکول اعتباری<sup>۴</sup> که ایجاد شدند به دنبال این بودند تا در حفظ ارزش اوراق با پشتوانه وام رهنی در برابر خطر تغییرات ناگهانی اطمینان حاصل کنند. فرض بر این بود که از پیچیدگی ریاضی در راستای کاهش خطر (ریسک) استفاده

1. Countrywide

بزرگ‌ترین گروه مالی در بریتانیا در حوزه املاک. م

2. Lehman Brothers

3. Moody

4. Credit Default Swaps

شود اما در عوض این امر منجر به ناتوانی تعداد محدودی تحلیلگر در درک روشن از ماوقع شد و همچنین ایجاد ابزارهای مالی محرمانه نظارت موثر را توسط روسای داخلی و ناظران خارجی تقریباً غیرممکن ساخت. به این جادوگران مشکوک، معیارهای مشکوک را نیز به عنوان جایگزینی برای قضاوت اضافه کنید؛ واقعیت این است که پاداش کارمندان برتر در بانک برادران لمان بر اساس پرداخت مبتنی بر عملکرد اندازه‌گیری شده به شکل تشویقی انجام می‌شد. بنابراین معیارها ابزارهایی ارائه می‌دهند و پرداخت مبتنی بر عملکرد انگیزه‌ها را برای کاهش ریسک‌پذیری در شرایط مبهم فراهم می‌کند. [۱۹] سپس از آنجا که وام‌گیرندگان ثابت کردند که نمی‌توانند وام‌های رهنی خود را پرداخت کنند و همچنین افت همزمان ارزش اوراق با پشتوانه وام رهنی منجر به خسارت‌های بسیار و غیر قابل پیش‌بینی به آن شرکت‌های مالی شد که اوراق بهادار را با تبادل افول اعتبار بیمه کرده بودند. نتیجه، فروپاشی تقریبی سیستم مالی بود.

### کوتاه‌مدت‌گرایی<sup>۱</sup>

روش دیگری که شاخص‌های مشکوک عملکرد اندازه‌گیری شده اقتصاد را تحریف کرده‌اند، کوتاه‌مدت‌گرایی است. شاید مهم‌ترین تغییر در دنیای کسب‌وکار و تجارت در دهه‌های اخیر مالی شدن اقتصاد بیشتر از همه در ایالات متحده آمریکا است. [۲۰] در اواخر دهه ۱۹۸۰، مالیه یک عنصر اصلی اما محدود اقتصاد آمریکا بود. مبادله سهام (بازار بورس) متشکل از سرمایه‌گذاران فردی بزرگ و کوچک بود که پول‌شان را روی سهام شرکت‌هایی سرمایه‌گذاری می‌کردند که باور داشتند در بلندمدت چشم‌انداز روشنی دارند. سرمایه‌گذاری از بانک‌های وال استریت (و دیگر همتایان خارجی) عرضه شده بود و در شرکت‌های خصوصی نیز پول شخصی شرکاء سرمایه‌گذاری می‌شد. همه اینها در راستای تبدیل به ذخایر بزرگ‌تر سرمایه (صندوق‌های بازنشستگی، موقوفات دانشگاه

1. Short-Termism

و سرمایه‌گذاران خارجی) آغاز به تغییر کردند و برای سرمایه‌گذاری در دسترس قرار گرفتند و به جای اینکه در دست خود صاحبان سرمایه باشند به دست مدیران حرفه‌ای پول افتادند. نتیجه، یک نظام مالی جدید بود که هایمن مینسکی<sup>۱</sup> آن را «سرمایه‌داری مدیر پول<sup>۲</sup>» و آلفرد راپاپورت<sup>۳</sup> استاد مدرسه کسب‌وکار آن را «سرمایه‌داری عاملیت<sup>۴</sup>» معرفی کردند. [۲۱]

وال استریت سنتی بخشی از این فرصت‌های جدید را تسریع کرد، بانک‌های سرمایه‌گذاری خود را به شرکت‌های بازرگانی عمومی تبدیل کردند - یعنی آنها نه تنها با پول خود که با پول دیگر افراد شروع به سرمایه‌گذاری کردند- و پاداش شرکا و کارمندان را به سود سالانه گره زدند. همه اینها یک نظام مالی بسیار رقابتی ایجاد کرد که توسط مدیران سرمایه‌گذاری با ذخایر هنگفت سرمایه اداره می‌شد و برای توانایی فرضی خود در عملکرد بهتر از همتایان خود «پاداش» می‌داد. ساختار مشوق‌های موجود در این محیط موجب شد مدیران تلاش کنند تا بازده‌های کوتاه‌مدت را به حداکثر برسانند و به نوبه خود این فشار را بر مدیران شرکت‌هایی آورد تا سود خود را هر سه ماه ارائه کنند. [۲۲]

افق زمانی کوتاه موجب ایجاد یک وسوسه در افزایش سود فوری با هزینه کردن افزایش سرمایه‌گذاری بلندمدت می‌شود؛ چه در تحقیق و توسعه و چه در افزایش مهارت‌های نیروی کار شرکت. تاکید بر سودهای سه ماهه (که قرار است موجب ارائه شفافیت شود) و «راهنمای سود سه ماهه» - طرح‌های سودآوری شرکت ظرف سه ماه - کوتاه‌مدت‌گرایی را تشدید می‌کند و آنها را همسو با معیارها تنزل می‌دهد و از آنجا که دست نیافتن به این هدف پیش‌بینی شده تا پایان سه‌ماهه بعدی ممکن است منجر به کاهش قیمت سهام شود، یک وسوسه غیر

1. Hyman Minsky

2. Money manager capitalism

3. Alfred Rappaport

4. Agency capitalism

قابل اجتناب برای بازی با اعداد و ارقام به وجود می‌آید تا عملکرد اندازه‌گیری‌شده با طرح‌ها مطابقت داشته باشد. همین موضوع انگیزه زیادی برای مدیران شرکت‌ها ایجاد می‌کند تا انرژی خلاقانه خود را به برنامه‌هایی اختصاص دهند که با دستکاری داده‌ها، بهره‌وری و سود خود را نشان یا با سرمایه‌گذاری کم روی تشکیل و حفظ سرمایه انسانی (آموزش مداوم کارمندان) سود سه‌ماهه یا معادل آن را افزایش دهند. گرایش به سرمایه‌گذاری در بلندمدت آنقدر به امری ناخوشایند بدل شد که در اوایل سال ۲۰۱۶، لری فینک<sup>۱</sup> مدیرعامل بزرگ‌ترین شرکت سرمایه‌گذاری، بلک‌راک<sup>۲</sup>، در نامه‌ای سرگشاده هشدار داد: «امروزه فرهنگ غیر قابل کنترل و هیجان‌آلود سود سه‌ماهه کاملاً با رویکرد بلندمدت که ما به آن نیاز داریم مغایر است.» [۲۳]

بازی معیارها اغلب موجب انحراف منابع از منافع بلندمدت به سوی اهداف کوتاه‌مدت می‌شود. شرکتی را در نظر بگیرید که با امید کسب درآمدهای چندگانه تلاش می‌کند که منافع خود را با اخراج نیروهای کار ضروری افزایش دهد. مدیرعاملی که با به تعویق انداختن سرمایه‌گذاری‌های مورد نیاز در تلاش برای برآورده ساختن انتظارات سه‌ماهه تحلیل‌گران، درآمد شرکت را ثابت نگاه دارد یا مدیرانی که در زمان گزارش‌های سه‌ماهه سهام خوب را می‌خرند و سهام بد را می‌فروشند؛ آنها در واقع این واقعیت را پنهان می‌کنند که سهام‌های خوب را با قیمت بالا خریده‌اند و سهام‌های بد فروخته‌شده نیز دوباره قیمت‌شان بالا می‌رود؛ این کار در تجارت به «ظاهرنمایی<sup>۳</sup>» شناخته می‌شود. [۲۴]

همان‌گونه که محققان سازمانی مانند نلسون رپنینگ<sup>۴</sup> و ربکا هندرسون<sup>۵</sup> اخیراً یادآور شدند؛ تمرکز بر شاخص‌های عملکرد اندازه‌گیری‌شده می‌تواند منجر به این شود که مدیران از وظایف خود که اندازه‌های

1. Larry Fink
2. BlackRock
3. Window Dressing
4. Nelson Repenning
5. Rebecca Henderson

روشنی از عملکرد ندارند غافل شوند [۲۵]. شمارش دارایی‌های نامشهود مانند اعتبار، رضایت شغلی، انگیزه، وفاداری، اعتماد و همکاری ممکن نیست و افرادی که طرفدار معیارهای عملکرد هستند، دارایی‌ها را در کوتاه‌مدت با هزینه نتایج بلندمدت افزایش می‌دهند. به همین دلیل، اتکا به معیارهای قابل اندازه‌گیری به کوتاه‌مدت‌گرایی ختم می‌شود؛ طاعون شرکت‌های آمریکایی معاصر.

### کژکارکردی‌های دیگر

هنگامی که پاداش‌ها مانند پرداخت، جایزه و ارتقا با اهداف بودجه گره می‌خورد، خطر دیگری نیز وجود دارد: تحریف سیستم اطلاعاتی سازمان. مدیران و کارمندان یاد می‌گیرند که دروغ بگویند، دستکاری کنند، پیرایش کنند و یا اعدادی را که برای محاسبه پرداختشان مورد استفاده قرار می‌گیرد، تغییر دهند. اما از آنجا که تعداد اعدادی که مدیران از آنها برای هماهنگی فعالیت‌ها و تصمیم‌گیری برای تخصیص منابع در آینده مورد استفاده قرار می‌دهند بسیار زیاد است، کارایی و بهره‌وری سازمان در صورت تخصیص اشتباه منابع آسیب می‌بیند. [۲۶]

همچنین تلاش برای جایگزینی اندازه‌گیری دقیق برای قضاوت آگاهانه، نوآوری را نیز محدود می‌کند که لزوماً منجر به تخمین و خطر (ریسک) می‌شود. همان‌طور که دو استاد مدرسه کسب‌وکار گری پیسانو<sup>۱</sup> و ویلی شی<sup>۲</sup> استدلال کرده‌اند:

«بیشتر شرکت‌ها برای ارزیابی فرصت‌های سرمایه‌گذاری به روش‌های بسیار تحلیلی متوسل شده‌اند. با این حال، ارزیابی برنامه‌های بلندمدت تحقیق و توسعه با تکنیک‌های کمی بسیار دشوار است ... معمولاً داده‌ها یا حتی تخمین‌های معمول هم به‌سادگی در دسترس نیست. با این وجود، در نهایت این ابزارها اغلب به داور نهایی درباره اینکه چه چیزی تامین مالی شود و چه چیزی نشود تبدیل می‌شود. از این‌رو، پروژه‌های کوتاه‌مدت با خروجی‌های قابل پیش‌بینی‌تر بر سرمایه‌گذاری‌های

1. Gary Pisano

2. Willy Shih

بلندمدت که نیاز به ذخیره‌سازی ظرفیت‌های عملیاتی و فنی دارند، پیروز می‌شود.» [۲۷]

معیارهای عملکرد به عنوان سنجشی برای پاسخگویی، زمانی که کارها خوب پیش نمی‌رود به شما کمک می‌کند تا خود را سرزنش کنید اما موفقیت را تشویق نمی‌کند [۲۸] به‌ویژه هنگامی که موفقیت به خیال‌پردازی، نوآوری و خطر کردن نیاز داشته باشد. در واقع، همانطور که اقتصاددان فرانک نایت حدود یک قرن پیش یادآور شده است؛ کارآفرینی مستلزم «عدم قطعیت بیکران» است که مستعد محاسبات معیاری نیست. [۲۹]

بنابراین، در کسب‌وکار و مالیه نیز ثبات معیارها عوارض خود را دارد. کسب‌وکار باید بر اساس بیش از یک شاخص عملکرد ارزیابی شود. مطمئناً سود مهم است اما در بلندمدت این اعتبار، سهم بازار، رضایت شغلی و روحیه کارکنان است که یافتن راه حل برای مشکلات جدیدی را که به طور گریزناپذیری در بازار به وجود می‌آید، ممکن می‌سازد. در دنیای اقتصادی که با تغییر غیر قابل پیش‌بینی، مشخص می‌شود نیاز به نوآوری مداوم در سطوح کوچک و بزرگ وجود دارد که نمی‌توان آن را به‌سادگی به یک هدف عملکرد تقلیل داد. مطمئناً شاخص‌های عملکرد می‌تواند کمک‌کننده باشد اما نمی‌تواند جایگزین عملکردهای کلیدی مدیریت شود: دوراندیشی، قضاوت، تصمیم‌گیری. [۳۰]

## فصل ۱۳

## بشر دوستی و کمک خارجی

همان‌طور که دیدیم، معیارهای عملکرد به‌سادگی موجب نقص و اختلال در موقعیت دولت و کسب‌وکارهای سودآور می‌شود و این امر در مورد سازمان‌های غیرانتفاعی نیز صادق است. مانند مطالعات موردی پیشین، هدف ما بررسی این حوزه به منزله یک کل نیست بلکه بررسی برخی مطالعات موردی و نمونه‌هاست.

خیریه‌ها نیز مانند شرکت‌ها و سازمان‌های دولتی تحت فشار هستند تا در کارهای خود شفاف و نسبت به خیرین خود پاسخگو باشند و اغلب تصور می‌شود که مطمئن‌ترین راه برای انجام این کار استفاده از معیارهای عملکرد است. گمان می‌رود که خیرین می‌خواهند کمک‌های آنها به‌طور موثری مورد استفاده قرار گیرد؛ مطابق با اهداف تعیین‌شده سازمان‌های غیرانتفاعی که در امور خیریه آنها مشارکت کرده‌اند. اما این امر چگونه ارزیابی می‌شود؟ و مشارکت‌کنندگان چگونه اطمینان حاصل کنند که وجود پرداختی آنها در درجه اول به نفع کارمندان خیریه مصرف نشده است؟

در دهه‌های اخیر، تحت طلسم ثبات معیارها، تامین‌کنندگان -بنیادهای تامین‌کننده سرمایه، دولت‌ها و افراد- تصمیم گرفتند که راه حل اندازه‌گیری

و عمومی کردن درصد بودجه هر خیریه مختص به هزینه‌های اداری و جذب سرمایه («هزینه‌های جاری» یا «غیرمستقیم») باشد مگر اینکه این امر با فعالیت‌ها یا برنامه‌های خیریه مغایرت داشته باشد. یک بار دیگر شاهد الگویی هستیم که ما را به استفاده از معیارها مواجه می‌کند. آنچه که اندازه‌گیری می‌شود در حقیقت آن چیزی است که به ساده‌ترین شکل ممکن اندازه‌گیری می‌شود و از آنجا که اندازه‌گیری خروجی‌های سازمان‌های خیریه دشوارتر از ورودی‌های آنهاست، این ورودی‌ها هستند که مورد توجه و بررسی قرار می‌گیرند. به شکل افراطی، نسبت هزینه‌های جاری به برنامه می‌تواند شاخص سودمندی برای ثقل یا مدیریت ضعیف مالی باشد. اما اغلب، عملکرد اندازه‌گیری شده‌ای که در موارد نابه‌جا مفید باشد به کلیه موارد گسترش می‌یابد.

برای بیشتر خیریه‌ها، برابر بودن هزینه‌های جاری (کلی) پایین با بهره‌وری بالاتر نه تنها فریبنده نیست بلکه کاملاً غیرمولد (آسیب‌زا) است. برای رسیدن به موفقیت، سازمان‌های خیریه نیاز به کارمندان مجرب و آموزش‌دیده دارند. آنها نیازمند کامپیوتر و سیستم‌های اطلاعاتی و دفاتر کاربردی نیز هستند و همچنین نیاز به تامین بودجه به طور مداوم دارند. اما این فرض که اثربخشی خیریه‌ها به طور معکوسی متناسب با هزینه‌های جاری‌شان است، منجر به کم کردن مخارج کلی و کاهش ظرفیت سازمانی می‌شود: به جای کارکنان مجرب و آموزش‌دیده، تعداد بسیار زیادی افراد تازه‌کار با برنامه‌گردش کارکنان مشغول به کارند و کامپیوترهای قدیمی و ناکارآمد مورد استفاده قرار می‌گیرد و در نتیجه این موارد موجب اثربخشی کمتر در تامین بودجه برای فعالیت‌های جاری یا برنامه‌های جدید می‌شود. تامین‌کنندگان نیز به طور فزاینده‌ای خواستار گزارش هستند، از این رو زمان کارکنان صرف مستندسازی و بخش بزرگی از بودجه صرف این امر می‌شود که اوضاع را بیش از پیش بدتر می‌کند. در واکنش به این موارد، روسای سازمان‌های خیریه غالباً تلاش می‌کنند تا با ارقام بازی کنند: با این گزارش که تقریباً بیشتر زمان اعضای پیشرو سازمان صرف این امور می‌شود یا اینکه هزینه‌ای برای تامین بودجه وجود ندارد. این واکنش قابل درک است اما همین واکنش‌ها انتظارات سرمایه‌گذاران را تامین می‌کند که هزینه‌های

جاری کمتر، معیاری است که آنها باید با توجه به آنها خیریه‌ها را پاسخگو نگاه دارند. [۱] بنابراین مارِ پاسخگویی، دُمِ خود را می‌خورد.<sup>۱</sup>

### تحول‌گرا در برابر قابل اندازه‌گیری

ثبات معیارها، شاهدهی است بر کمک خارجی دولت که هدف آن افزایش توسعه اجتماعی و اقتصادی است. یک شک و تردید عمیقاً ریشه‌دار و تا حدودی مستدل درباره کمک خارجی وجود دارد که اغلب آن را ناپه‌ره‌ور و در واقع ضدمولد می‌دانند. [۲] اما بعضی از برنامه‌های کمک خارجی حقیقتاً به سلامت، آموزش، توسعه اقتصادی و حتی ثبات سیاسی کشورهای فقیر کمک می‌کنند. در تلاش برای سنجش آنچه که موثر است و آنچه که نیست، سازمان‌های دولتی آمریکا به طور فزاینده‌ای به معیارها نظر داشتند؛ با نتایجی که می‌تواند برای خوانندگان این کتاب قابل پیش‌بینی باشد.

برنامه‌هایی که اندازه‌گیری دستاوردهای آنها به شکل کمی ساده نیست بسیار محدود شده‌اند. مثلاً، سنجش ثبت‌نام در مدارس ابتدایی و نرخ سواد ساده‌تر از آموزش فرهنگی نخبگان آینده است که از طریق بورسیه‌های دانش‌آموزی کشورهای فقیر برای تحصیل در دانشگاه‌های آمریکا ارائه می‌شود. بنابراین وقتی معیارها به استاندارد ارزیابی تبدیل می‌شود، برنامه‌هایی که نمی‌تواند مزایای کوتاه‌مدت خود را نشان دهد، قربانی می‌شود. مثلاً برنامه بورسیه تحصیلی آژانس توسعه بین‌المللی ایالات متحده آمریکا توسط دفتر مدیریت و بودجه کاخ سفید حذف شد به این دلیل که مزایای آن به صورت دلاری قابل بررسی نیست و بنابراین دولت نمی‌تواند بفهمد که آیا مزایای این برنامه از هزینه‌های آن بیشتر است یا نه. [۳]

در اینجا نیز معیارها کوتاه‌مدت‌گرایی را ترویج می‌دهد. اندرو ناتسیوس<sup>۲</sup>، کارمند برجسته دولتی با تجربه طولانی در توسعه بین‌المللی خاطر نشان می‌کند که کارمندان سازمان‌های دولتی در این زمینه «مبتلا به یک مورد بسیار بد اختلال اندازه‌گیری و سواس‌آلود شده‌اند؛ یک اختلال فکری که ریشه در این مفهوم دارد که

۱. اشاره به نماد باستانی اوروبوروس که نماد کیمیاگری دنیای باستان است و ریشه‌های آن به مصر باستان بازمی‌گردد. این نماد یک مار است که در حال خوردن دم خود است. این نماد نشان‌دهنده ماهیت چرخشی زندگی و بازگشت ابدی و زندگی جاودان و چرخه ابدی است که امروزه به عنوان تناسخ شناخته می‌شود. م

2. Andrew Natsios

شمارش همه چیز در برنامه‌های دولت منجر به انتخاب‌های سیاسی بهتر و بهبود مدیریت می‌شود.» تاکید بر کمیت منجر به غفلت از برنامه‌هایی با مزایای بالقوه بلندمدت می‌شود: مواردی که مهارت‌ها، دانش و هنجارهای خدمات اجتماعی و سیستم‌های قضایی در کشورهای توسعه‌نیافته را بهبود می‌دهد. ناتسیوس می‌نویسد کسانی که از اختلال اندازه‌گیری وسواس‌آلود رنج می‌برند «یک اصل کلیدی نظریه توسعه را نادیده می‌گیرند؛ اصلی که بر مبنای آن برنامه‌های توسعه‌ای که سریع و ساده اندازه‌گیری می‌شوند کمترین میزان تحول‌گرایی را دارند و برنامه‌هایی که بیشترین میزان تحول‌گرایی را دارند، قابلیت اندازه‌گیری‌شان حداقل است.» [۴]

البته در این میان روسا و مدیران مجرب نیز رشد و پیشرفت خوبی دارند.

در اینجا پافشاری بر ساده‌ترین موارد قابل اندازه‌گیری منجر به تمرکز نه تنها بر روی خروجی‌ها بلکه بر ورودی‌های قابل اندازه‌گیری مانند فرآیندهای دیوان‌سالاری شده است. همان‌گونه که یک مقام آژانس توسعه بین‌المللی ایالات متحده آمریکا به یک پژوهشگر اعتراف کرده است: «هیچ‌کس روش معتبری برای تعیین اثربخشی فعالیت‌های ایجاد ظرفیت ندارد ... از این رو به جای تمرکز بر اثربخشی در گزارش، آژانس توسعه بین‌المللی آمریکا بر آنچه که می‌توان اندازه‌گیری کرد تمرکز دارد، مانند تعداد کارگاه‌هایی که برگزار می‌شود یا تعداد افرادی که در دوره آموزشی شرکت می‌کنند.» [۵]

تقاضا برای اندازه‌گیری و سنجش بیشتر نه تنها از کمیته‌های کنگره نشأت می‌گیرد، بلکه از سوی دستگاه‌های اجرایی مانند دفتر مدیریت و بودجه و دفتر پاسخگویی دولتی و برخی سازمان‌های دیگر درخواست می‌شود؛ سازمان‌هایی که دارای کارکنانی خوب مانند «حسابداران، اقتصاددانان، مقامات تامین‌کننده و قانون‌گذار» ... و دارای مهر تایید استادانی در حوزه‌های مدیریت دولتی، مدیریت کسب‌وکار یا اقتصاد هستند که تاکید بیش از اندازه‌ای بر کمیت و اندازه‌گیری در آموزش آنها داشته‌اند. [۶] این اندازه‌گیران حرفه‌ای، راهبه‌های پاکدامن آتش مقدس معیارها هستند. آنها همچون مبلغان مذهبی هستند که مدیریت ارشد خود را به یک مکتب فکری و آیین دینی بدل می‌کنند که خواستار تخصیص قابل توجه زمان و انرژی به ارائه گزارش‌های آماری است که با استفاده از آنها بتوان عملکرد را اندازه‌گیری کرد و از وجود پاسخگویی مطمئن بود.

ضمیمہ



## فصل ۱۴

## هنگامی که شفافیت دشمن عملکرد است

### سیاست، دیپلماسی، اطلاعات و ازدواج

جذابیت معیارها در بخش خوب این مفهوم نهفته است که نهادها اگر غیرشفاف باشند، غیرمسئول و غیرپاسخگو نیز هستند و اگر تحت کنترل و نظارت بیرونی باشند موثرتر خواهند بود. گوگل انگرام نشان می‌دهد که در اواسط دهه ۱۹۸۰ جست‌وجوی دو اصطلاح «معیارهای عملکرد» و «شفافیت» همراه با یک شیب تند افزایش داشته است. ویژگی فرهنگ ما گرایش به این فرض است که عملکرد و شفافیت در کنار هم افزایش و کاهش می‌یابد. اما این یک مغلطه یا دست‌کم یک تعمیم گمراه‌کننده است. همان‌طور که محدودیت‌هایی در کارایی عملکرد اندازه‌گیری شده وجود دارد، محدودیت‌هایی نیز برای کارایی شفافیت وجود دارد. در برخی موارد، عملکرد نهادهای ما وابسته به شفاف بودن نیست. در اینجا مساله معیارها نیست، بلکه عملکرد در وسیع‌ترین گستره ممکن است: مساله، موفقیت در کاری است که قرار است انجام دهیم. برای ارزش نهادن بر بخش شفافیت، اجازه دهید که نه با سازمان‌ها که با روابط فردی آغاز کنیم.

## صمیمیت

خویشتن‌پنداری (حس واقعی ما از خودمان) تنها به این دلیل ممکن است که تفکرات و خواست‌های ما برای دیگران شفاف نیست. ایجاد صمیمیت به توانایی ما در میزان شفاف بودن‌مان در ارتباط با دیگران بستگی دارد. همان‌طور که فیلسوف معاصر موشه هالبرتال<sup>۱</sup> می‌گوید: «اگر افکار یک شخص روی پیشانیش نوشته شده باشد، پیش از هر چیز تمایز میان درون و بیرون و البته فردیت از بین می‌رود. حریم خصوصی از طریق پنهان‌سازی بروز می‌کند. بنابراین حفظ حریم توانایی یک فرد در تعریف فردیت خود است. علاوه بر این، خود<sup>۲</sup> ممکن است با نشان دادن ابعاد مختلف آسیب‌پذیری و صمیمیت روابط ویژه‌ای ایجاد کند. او از فضای اجتماعی با تخصیص آشکارسازی و پنهان شدن و ابعاد مختلف فاصله و نزدیکی عبور می‌کند.» [۱]

موفقیت روابط فردی حتی در صمیمی‌ترین آنها وابسته به میزانی از ابهام و نبود شفافیت است و اینکه هرگز به همه کارهایی که دیگری انجام می‌دهد نباید فکر کرد.

## سیاست و دولت

وقتی صحبت از سیاست می‌شود -جایی که بازیگران بسیاری حضور دارند و از این رو حساسیت‌ها و منافع بیشتری نیز وجود دارد- میزان مشخصی از ابهام بیش از پیش مورد نیاز است. یکی از مهم‌ترین نقش‌های سیاست‌مداران پادرمیانی میان منافع و حساسیت‌ها و ایجاد پُل میان تفاوت‌هاست. این استراتژی مستلزم مذاکره است؛ چشم‌پوشی از برخی منافع در برابر منافع دیگر در راستای ایجاد توافقی که همسو با منافع تعداد بیشتری از افراد است، هرچند در نهایت به ندرت می‌توان همه را راضی نگاه داشت. به عبارت دیگر، این امر مستلزم چانه‌زنی در بسیاری از موقعیت‌هاست؛ حداقل در دفاع از طرف‌های ذی‌نفع. اکثر اوقات این امر ممکن نیست یا تنها در صورتی ممکن است که مذاکرات

1. Moshe Halbertal

2. The self

از دید مطالبه‌گران و مدعیان دور باشد، تا آنجا که هر یک از طرف‌های مذاکره‌کننده تلاش می‌کنند توافقی را که ادعای «شفافیت» دارد، وتو کنند. آنچه که سیاست‌مداران آن را «بده بستان خلاق» می‌نامند، ایدئولوژیست‌ها یا نمایندگان‌شان که دارای منافع ویژه‌ای هستند «خیانت» می‌دانند. به همین دلیل فرآیند مذاکره در موضوعات حساس زمانی موثرتر است که پشت درهای بسته انجام شود. همان‌گونه که تام دشله<sup>۱</sup> رهبر پیشین اکثریت دموکرات سنا مشاهده کرده است: «این ایده که با وجود دوربین‌های مداربسته تلویزیونی که همه مکالمات را ضبط و کنترل می‌کنند، واشنگتن بهتر کار خواهد کرد کاملاً اشتباه است ... نبود فرصت برای گفت‌وگوی صادقانه و بده بستان خلاق ریشه در همین نقص دارد.» [۲] به همین دلیل است که سیاست‌مداران موثر باید تا حدودی از هم فاصله داشته باشند و با درجه‌ای از انعطاف‌پذیری مذاکرات را پشت درهای بسته به جای عرصه عمومی دنبال کنند. تنها هنگامی که توافقی‌های مختلف و معاملاتی حاصل شده است باید تحت نظارت عمومی قرار گیرد؛ یعنی شفاف‌سازی شود. [۳] همین امر در مورد عملکرد دولت نیز صادق است؛ در اینجا نیز عملکرد موثر به باز نبودن درهای جلسات داخلی بر روی عموم وابسته است - یعنی حفظ نبود شفافیت. ما نیاز داریم که میان عناصر دولتی که باید علنی شود و آنها که نباید علنی شود، تمایز قائل شویم. کاس آر. سانستین<sup>۲</sup> پژوهشگری که در دولت نیز خدمت کرده است تمایز مفیدی میان ورودی‌ها و خروجی‌های دولت قائل است. خروجی‌ها شامل داده‌های تولید شده توسط دولت در حوزه‌های اجتماعی و اقتصادی و نتایج اقدامات دولت مانند قوانین نظارتی هستند. او استدلال می‌کند که خروجی‌ها باید تا حد ممکن در دسترس عموم قرار گیرند. ورودی‌ها برعکس خروجی‌ها بحث‌هایی است که برای تصمیم‌گیری به دولت می‌روند؛ بحث میان سیاست‌گذاران و کارمندان دولت. فشارهای فزاینده‌ای برای ایجاد دسترسی عمومی وجود دارد؛ خواه از طریق ابزارهای قانونی باشد مانند

1. Tom Daschle

2. Cass R. Sunstein

درخواست‌های قانون آزادی اطلاعات یا خواسته‌های کنگره مانند موردی که کمیته‌های کنگره خواستار مکاتبات ایمیلی وزیر امور خارجه پیشین هیلاری کلینتون در پرونده تحقیقات بنگازی شد یا ابزارهای غیرقانونی مانند سرقت الکترونیکی و انتشار اسناد داخلی دولت توسط سازمان‌ها مانند ویکی‌لیکس. باز کردن درهای جلسات داخلی به روی عموم - که در حکم شفافیت شناخته می‌شود - غیرمولد است، سانس‌تین استدلال می‌کند که اگر مقامات دولتی بدانند همه ایده‌ها و مواضع‌شان ممکن است به طور عمومی فاش شود، در ارتباطات‌شان از شفافیت، صداقت و اعتماد پرهیز می‌کنند. نتیجه قابل پیش‌بینی این خواهد بود که مقامات دولتی هر روز مطالب کمتری می‌نویسند چه به صورت چاپی و چه به شکل ایمیل. در عوض آنها موضوعات مهم را به مکالمات شفاهی محدود می‌کنند. اما همین موضوع فرصت را برای تعیین دقیق مواضع کاهش می‌دهد. [۴] همه سیاست‌ها هزینه‌بر هستند، اگر جلسات داخلی به روی عموم باز باشد غیرممکن است که بتوان از تجویزهای سیاسی که محبوب اما غیرعاقلانه باشند جلوگیری یا سیاست‌هایی را رد کرد که مطلوبند، اما موجب آسیب و رنجش دیگر رای‌دهندگان می‌شود. بنابراین شفافیت ورودی‌ها تبدیل به دشمن دولت خوب می‌شود.

### دیپلماسی و جمع‌آوری اطلاعات

شفافیت یک خطر در دیپلماسی و برای جمع‌آوری اطلاعات کشنده است. در سال ۲۰۱۰ برَدلی مانینگ<sup>۱</sup> که تحلیلگر اطلاعاتی در ارتش آمریکا بود پذیرفت که صدها هزار اسناد مربوط به وزارت خارجه و ارتش را از طریق ویکی‌لیکس افشا کرده است. [۵] نتیجه این امر انتشار نام شاهدان به صورت محرمانه بود؛ شامل دگراندیشان سیاسی که با دیپلمات‌های آمریکایی در خاورمیانه، چین، افغانستان، جهان عرب و جاهای دیگر صحبت کرده بودند. [۶] در نتیجه، برخی از این افراد برای حفظ جان خود مجبور به جابه‌جایی شدند. مهم‌تر از همه، این

1. Bradley Manning

افشاگری‌ها دستیابی به اطلاعات انسان در آینده را برای دیپلمات‌های آمریکایی بسیار دشوار کرده است زیرا دیگر نمی‌توان از محرمانه بودن مکالمات اطمینان داشت.

سپس در سال ۲۰۱۳ ادوارد اسنودن<sup>۱</sup> که پیشتر متخصص امنیت کامپیوتر در سیا و سپس پیمانکار آژانس امنیت ملی در هاوایی بود به طور سیستماتیک اقدام به رونوشت از هزاران اسناد کاملاً محرمانه از سازمان‌های مختلف دولتی به منظور افشای برنامه‌های جاسوسی دولت آمریکا کرد. در میان بسیاری از اسناد حساسی که در اختیار مطبوعات قرار گرفت یک سند ۱۸ صفحه‌ای بود؛ دستورالعمل سیاسی ریاست جمهوری ۲۰<sup>۲</sup> در مورد عملیات سایبری. این سند افشا می‌کرد که هر سیستم کامپیوتر خارجی برای اقدام بالقوه نیز مورد هدف قرار گرفته است - سندی که به طور کامل توسط روزنامه بریتانیایی گاردین<sup>۳</sup> منتشر شد. افشای اسنادی که توسط اسنودن به سرعت رفته و توسط رسانه‌های پیشرو منتشر شد، نه تنها نشان از مهم‌ترین رخنه در سرویس‌های اطلاعاتی آمریکا داشت بلکه نشان‌دهنده یک ضربه قدرتمند به امنیت ملی ایالات متحده آمریکا و دوستان و متحدانش بود. با این حال از اسنودن در بخش‌هایی از ایالات متحده و اروپا به عنوان یک قهرمان استقبال شد. در قلب فاجعه اسنودن این باور نهفته وجود دارد که شفافیت همیشه مطلوب است.

یک خط‌مشی موفق مانند ازدواج سالم، برخی موضوعات را به سایه می‌راند. در روابط بین‌الملل نیز مانند روابط فردی بسیاری از اقدامات تا زمانی که مبهم و غیرشفاف هستند کاربرد دارند. شفافیت و افشاگری کشنده‌اند. توانایی مذاکره میان زوجین یا دولت‌ها اغلب در بردارنده فرمول‌هایی است که به هر یک از طرفین اجازه می‌دهد چهره و شخصیت خود را حفظ کنند و این امر یا نیازمند توافق است یا ابهام. واقعیت این است که متحدان برای تعیین اهداف، ظرفیت‌ها و آسیب‌پذیری‌هایی که برای شاغلان دولتی به خوبی شناخته شده است از یکدیگر جاسوسی می‌کنند. اما نمی‌توان این امر را به طور عمومی

1. Edward Snowden

2. Presidential Policy Directive 20 (PPD-20)

3. The Guardian

تأیید کرد زیرا عزت نفس دیگر ملت‌ها را تهدید می‌کند. علاوه بر این، در سیاست داخلی و روابط بین‌الملل همچون روابط فردی، میزانی از عدم خلوص (دورویی) برای اقدامات تساهل‌محور و سودمند نقش دارد. هر چند با قوانین بین‌المللی و هنجارهای صریح قابل توجیه نیستند. به طور خلاصه، نقل قولی از موشه هالبرتال را در همین راستا داریم:

«برای حفظ دولت و نهادهای دموکراتیک آن، درجه‌ای از پنهان‌کاری مشروع لازم است. اسرار نظامی، تکنیک‌های مبارزه با جرم، جمع‌آوری اطلاعات و حتی مذاکرات دیپلماتیک که اگر افشا شود، از بین خواهد رفت - همه این حوزه‌ها باید در راستای امکان شفافیت معمول در دیگر نهادهای دولت در خفا پنهان بمانند. گفت‌وگوی آزاد و شفاف ما بیشتر وابسته به حوزه پنهان و بسیار تاریکی است که موفقیت آن را تضمین می‌کند.» [۷]

ما در دنیایی زندگی می‌کنیم که حریم خصوصی هم از طریق فناوری (اینترنت) و هم از طریق فرهنگی در حال از بین رفتن است که فضیلت عدم محدودیت را بدون نیاز به شرم اعلام می‌کند. در چنین جامعه‌ای که حریم خصوصی را پشت سر گذاشته است، مردم می‌خواهند از ارزش پنهان بودن چشم‌پوشی کنند. [۸] بنابراین قدرت «شفافیت» به عنوان یک فرمول جادویی است که اثرات مخرب آن اغلب نادیده گرفته می‌شود. «نور خورشید بهترین ضد عفونی کننده است.» تبدیل به بیانیه اعتقادی ایمان جدید به ویکی‌لیکس‌سم شده است: این عقیده که افشاسازی تاملات و تفکرات در همه سازمان‌ها و دولت‌ها، جهان را به مکانی بهتر تبدیل می‌کند.

اما نتیجه اغلب فلج‌کننده است. سیاست‌مدارانی که مجبورند هر اقدام خود را عمومی کنند، قادر به دستیابی به توافق‌هایی نیستند که قانون‌گذاری را ممکن می‌سازد. مقاماتی که از آشکار شدن فعالیت‌ها و اندیشه‌های خود واهمه دارند، توانایی اثرگذاری کمتری بر سیاست‌گذاری عمومی خواهند داشت. سازمان‌های اطلاعاتی که نیازمند جمع‌آوری اطلاعات از دشمنان خود به طور مخفیانه هستند، نیز دچار مشکل شده‌اند. در هر حالت، شفافیت به دشمن عملکرد تبدیل شده است.

## نتیجه گیری



## فصل ۱۵

# پیامدهای منفی ناخواسته اما قابل پیش‌بینی

یکی از مفاهیم مهم علوم اجتماعی در قرن نوزدهم توسط آگوست کامت<sup>۱</sup> ابداع شده است: بدانید برای پیش‌بینی کردن، پیش‌بینی کنید برای پیش‌گیری کردن از پیامدهای پیش‌بینی‌نشده اقدامات<sup>۲</sup>. اکنون که ما دانش خوبی درباره ثبات معیارها داریم، می‌توانیم بسیاری از پیامدهای منفی ناخواسته آن را پیش‌بینی و شاید از بروز آنها جلوگیری کنیم. پیش از آنکه از عملکرد اندازه‌گیری‌شده استفاده درستی کنیم، اجازه دهید چند درس آموخته از مطالعات موردی درباره خطر تکراری معیارها را مرور کنیم.

**جابه‌جایی هدف از طریق انحراف تلاش به سوی آنچه که اندازه‌گیری می‌شود:** جابه‌جایی اهداف، انواع مختلفی دارد. هنگامی که درباره عملکرد با چند معیار قضاوت می‌شود و خطرات بالاست (نگه داشتن شغل شخص، بالا بردن رده شغلی، افزایش قیمت سهام هنگامی که اختیارات

1. Auguste Comte

2. Savoir pour prévoir, prévoir pour prévenir

سهام واگذار شده است) - افراد بر همین معیارها تمرکز می‌کنند؛ اغلب هم به هزینه بی‌توجهی به اهداف سازمانی مهم‌تر دیگری که اندازه‌گیری نمی‌شود. [۱] اقتصاددانان، بنت هولمستروم<sup>۱</sup> و پُل میلگرام<sup>۲</sup> آن را به صورت رسمی‌تر به عنوان مشکل انگیزه‌های نامرتب<sup>۳</sup> تعریف کرده‌اند: کارکنانی که برای انجام وظایف قابل اندازه‌گیری پاداش می‌گیرند، از کوشش برای انجام دیگر وظایف می‌کاهند. [۲] نتیجه این می‌شود که معیارها به جای اینکه ابزارهایی برای خدمت‌رسانی باشند، خودشان به اهداف سازمانی بدل می‌شوند.

**ترویج کوتاه‌مدت‌گرایی:** عملکرد اندازه‌گیری شده موجب تشویق آن چیزی می‌شود که رابرت کی. مِرتون<sup>۴</sup> آن را اینچنین توصیف می‌کند: «منافع فوری بی‌واسطه... جایی که دغدغه اصلی بازیگران با عواقب آنی پیش‌بینی شده مانع توجه به آینده و پیامدهای آن می‌شود.» [۳] به طور خلاصه، هزینه پیشبرد اهداف کوتاه‌مدت، ملاحظات بلندمدت است.

**هزینه‌های زمان شاغلی<sup>۵</sup>:** به سمت بدهکار دفتر کل باید هزینه‌های مبادله معیارها را نیز افزود: هزینه زمان شاغلی توسط کسانی که وظیفه تهیه و پردازش معیارها را دارند؛ صحبت از زمان مورد نیاز برای خواندن معیارها نیست. این امر با «ضرورت گزارش‌دهی» تشدید می‌شود؛ نیاز به تولید مداوم اطلاعات، حتی زمانی که هیچ چیز قابل توجهی در جریان نیست. گاهی معیار موفقیت، تعداد و میزان گزارش‌های تولید شده است؛ گویی هیچ کاری محقق نمی‌شود مگر اینکه به طور گسترده مستند شود. افراد درون سازمان‌ها زمان بیشتر و بیشتری را صرف جمع‌آوری داده، نوشتن گزارش و شرکت در جلسه‌هایی می‌کنند که هدف از برگزاری آنها نیز طبقه‌بندی و تنظیم داده‌ها و گزارش‌هاست. بنابراین، همان‌گونه که مشاوران دگراندیش مدیریت ایو موریو<sup>۶</sup> و پیتر تولمن<sup>۷</sup> خاطرنشان می‌کنند، کارمندان در فعالیت‌هایی که

1. Bengt Holmström

2. Paul Milgrom

3. Misaligned

4. Robert K. Merton

5. Costs in employee time

6. Yves Morieux

7. Peter Tollman

حقیقتاً بهره‌وری اندکی برای سازمان‌شان دارند و در حالی که موجب از بین رفتن شور و اشتیاق آنان می‌شود، بسیار سخت و طولانی کار می‌کنند. [۴]

**کاهش مطلوبیت:** گاهی معیارهای عملکرد که به‌تازگی معرفی شده‌اند مزایایی فوری در کشف داده‌های عملکرد خارج از محدوده و ضعیف دارند. [۵] با وجود انجام ساده‌ترین کارها، انتظار نتایج فوق‌العاده وجود دارد. مشکل این است که معیارها از همه افراد جمع‌آوری می‌شود و خیلی زود هزینه‌های گردآوری و تحلیل معیارها از مزایای حاشیه‌ای آنها فراتر می‌رود.

**جریان‌های (رویه‌های) قانونی:** سازمان‌ها در تلاش برای ایجاد جریان‌های ناقص از طریق بازی، تقلب و انحراف اهداف، جریان‌های (رویه‌های) قانونی ایجاد می‌کنند که اجرای آنها عملکرد سازمان را کندتر می‌کند و کارایی را کاهش می‌دهد.

**شانس پاداش گرفتن:** اندازه‌گیری خروجی‌ها زمانی که افراد کنترل اندکی بر نتایج دارند برابر است با شانس پاداش گرفتن. این بدان معناست که افراد به دلیل خروجی‌های مستقل از تلاش‌شان تشویق یا تنبیه می‌شوند. افراد تنبیه شده به‌درستی احساس می‌کنند که با آنها رفتار عادلانه‌ای صورت نگرفته است.

**دلسرد کردن ریسک‌پذیری:** تلاش برای اندازه‌گیری بهره‌وری از طریق معیارهای عملکرد اثرات هوشمندانه دیگری نیز دارد: آنها نه تنها کوتاه‌مدت‌گرایی را ترویج می‌دهند بلکه همان‌گونه که پیشتر گفتیم موجب دلسرد شدن ابتکار عمل و ریسک‌پذیری می‌شوند. تحلیلگران اطلاعاتی که سرانجام مکان بن‌لادن را یافتند، سال‌ها روی این مشکل کار می‌کردند. اگر اندازه‌گیری در هر نقطه‌ای از انجام این کار صورت می‌گرفت، به نظر می‌رسد که بهره‌وری آنان صفر باشد. تا ماه‌ها بعد شکست آنها صد درصد بود تا اینکه در نهایت موفق شدند. از دیدگاه مافوقان این تحلیلگران اطلاعاتی، دادن اجازه به تحلیلگران برای کار روی این پروژه به مدت چند سال، میزان ریسک‌پذیری آن را بسیار بالا می‌برد؛ سرمایه‌گذاری روی زمان ممکن است موفقیتی

در پی نداشته باشد. با این حال، دستاوردهای بزرگ اغلب به چنین ریسک‌هایی بستگی دارد. این امر لازمه موقعیت‌هایی است که درگیر سرمایه‌گذاری‌های بلندمدت روی نیروی انسانی است.

**دلسرد کردن نوآوری:** هنگامی که افراد بر اساس معیارهای عملکرد قضاوت می‌شوند، برای انجام آنچه که معیارها اندازه‌گیری می‌کند انگیزه پیدا می‌کنند و آنچه که معیارها اندازه می‌گیرد برخی اهداف را تعیین می‌کند. اما همین موضوع مانعی برای نوآوری است؛ بدین معنا که کاری هنوز انجام نشده یا برای آن تلاشی صورت نگرفته است. نوآوری شامل آزمودن‌های بسیار است. آزمون چیزهای جدید مستلزم خطر کردن، امکان و شاید احتمال شکست است. [۶] هنگامی که معیارهای عملکرد، موجب تشویق به خطر نکردن می‌شود، سهواً رکود را دامن می‌زند.

**دلسرد کردن همکاری و هدف مشترک:** پاداش دادن به افراد بر اساس عملکرد اندازه‌گیری‌شده حس هدف مشترک را کاهش می‌دهد مانند روابط اجتماعی که انگیزه‌های غیرقابل اندازه‌گیری برای همکاری و اثربخشی نهادی فراهم می‌کند. [۷] پاداش مبتنی بر عملکرد اندازه‌گیری‌شده موجب ترویج همکاری و رقابت می‌شود. اگر افراد یا واحدها به جای کمک، مساعدت و مشاوره به یکدیگر به مشوق‌ها پاسخ دهند، تمام تلاش خود را می‌کنند تا معیارهای خود را به حداکثر برسانند و معیارهای دیگران را یا نادیده بگیرند یا حتی به آنها آسیب برسانند. همان‌گونه که دونالد پرویک<sup>۱</sup>، اصلاح‌گر پیشرو در پزشکی خاطرنشان می‌کند:

«مدیرعامل یک بیمارستان، سیستم مدیریت سودمحور خود را این‌طور برای من تشریح کرد که در آن جوایز مدیران میانی وابسته به عملکرد بودجه محلی است. از او پرسیدم: آیا یکی از مدیران منابع را از بخش خود به بخش دیگر منتقل می‌کند اگر این کار او موجب کمک به کل سازمان شود؟ مدیر عامل پاسخ داد: بله، اگر دیوانه باشد.» [۸]

1. Donald Berwick

**تنزل کار:** وادار کردن افراد در یک سازمان برای متمرکز کردن تلاش‌های‌شان در یک دامنه محدود اندازه‌گیری و سنجش به تخریب و تنزل تجربه کاری می‌انجامد. ادmond فلیپس<sup>۱</sup>، اقتصاددان برنده نوبل در کتابش با عنوان «شکوفایی انبوه: چگونه نوآوری مردمی شغل، چالشو تغییر ایجاد می‌کند»<sup>۲</sup>، این ادعا را مطرح می‌کند که یکی از فضایل سرمایه‌داری توانایی آن در ارائه «تجربه محرک ذهنی، چالش مشکلات جدید برای حل شدن، فرصت امتحان کردن چیزهای جدید و ایجاد هیجان برای ورود به ناشناخته‌هاست.» [۹] در حقیقت احتمال همه این موارد در سرمایه‌داری وجود دارد. اما آنهایی که مشمول معیارهای عملکرد هستند، مجبورند روی اهداف محدودی که توسط دیگران به آنها تحمیل می‌شود و ممکن است درک درستی از آن نداشته باشند، متمرکز شوند. برای کارکنانی که تحت نظارت شدید قرار دارند، محرک‌های ذهنی کم‌رنگ می‌شود، آنها تصمیم می‌گیرند که نه به حل مشکلات و نه به چگونگی حل آنها بپردازند، آنها هیچ هیجانی برای ورود به ناشناخته‌ها ندارند زیرا ناشناخته‌ها فراتر از قابل اندازه‌گیری‌هاست. به طور خلاصه، عنصر کارآفرینی طبیعت انسانی - که فراتر از صاحبان شرکت‌ها است - ممکن است با ثبات معیارها از بین برود. [۱۰]

یک نتیجه، تشویق افراد با ابتکار عمل و تلاش بیشتر است که از جریان غالب در سازمان‌های بزرگ که مبتنی بر فرهنگ پاسخگویی عملکرد است، خارج شوند. آموزگاران از مدارس دولتی بیرون بیایند و وارد مدارس خصوصی و خصلتی شوند. مهندسان از شرکت‌های بزرگ به شرکت‌های کوچک بروند و کارمندان مجرب دولت تبدیل به مشاور شوند. یک عنصر سالم در این امر وجود دارد اما مطمئناً سازمان‌های بزرگ جامعه ما فقیرتر از آنند که بتوانند دست به ابتکار عمل و نوآوری بزنند. هر چه کار پر کردن جعبه‌هایی که با معیارها پُر و به آنها پاداش داده می‌شود بیشتر باشد، افرادی هم که فکر می‌کنند بیرون از جعبه‌ها هستند بیشتر رد می‌شوند.

1. Edmund Phelps

2. Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge, and Change

**هزینه‌های بهره‌وری:** اقتصاددانان متخصص در سنجش بهره‌وری اقتصادی گزارش می‌دهند که در سال‌های اخیر تنها افزایش در عامل کل بهره‌وری در اقتصاد آمریکا در صنایع تولیدکننده فناوری اطلاعات بوده است. [۱۱] سوالی که باید پرسیده شود این است که آیا فرهنگ معیارها - با هزینه‌های زمان شاغلی، روحیه، ابتکار عمل و ترویج کوتاه‌مدت‌گرایی - به رکود اقتصادی کمک کرده است؟

## فصل ۱۶

## چه زمانی و چگونه از معیارها استفاده شود؟

## فهرست

هیچ چیز ذاتاً آسیب‌پذیر درباره شمارش و اندازه‌گیری عملکرد انسان وجود ندارد. همه ما می‌خواهیم نتیجه‌گیری‌های گسترده‌ای بر مبنای تجارب محدود بگیریم و داده‌های اندازه‌گیری‌شده می‌توانند به منزله یک نقطه مقابل مفید برای احکام ذهنی عمل کنند. نوع اندازه‌گیری‌های مرتبط با این کتاب، مربوط به معیارهای عملکرد است که میزان موفقیت و شکست انسان را به صورت کمی تعیین می‌کنند. تقریباً در هر سازمان، معیارهای عملکرد قانونی وجود دارد. در مطالعات موردی، نمونه‌های بسیاری را دیدیم که در آنها معیارها سودمند و موثر بودند.

در اداره پلیس، از آمارهای کامپیوتری وقوع جرایم برای اهداف مناسب در راستای دانستن این استفاده می‌شود که بیشترین مشکلات در کجا قرار دارند و نیروهای پلیس باید در چه مکان‌هایی مستقر شوند. مشکلات زمانی ایجاد شد که مقامات با تهدید افراد رده‌پایین در سلسله مراتب به تنزل رتبه تلاش کردند تا از گزارش میزان جرم و جنایت بکاهند.

در دانشگاه‌ها ارزیابی‌های اعضای هیات علمی می‌تواند با داده‌های عددی و مطالب منتشرشده و تدریس بهتر شود. معیارها زمانی که به صورت مکانیکی تهیه و توسط افرادی اعمال شود که در جایگاه ارزیابی صحت و اهمیت داده‌ها نباشند، شکست می‌خورد.

در آموزش ابتدایی و متوسطه، می‌توان از آزمون‌های استاندارد برای آگاه‌سازی آموزگاران در مورد میزان یادگیری دانش‌آموزان در موضوعات خاص استفاده کرد. معلمان می‌توانند با همکاران خود مشورت کنند و روش‌ها و برنامه‌های درسی خود را با هم هماهنگ کنند. مشکل از جایی نشأت می‌گیرد که آزمون، بدل به بنیان اصلی آموزگاران و مدارس برای پاداش و تنبیه می‌شود.

در پزشکی، پروژه کی‌استون پیتز پرونوست<sup>۱</sup> نشان می‌دهد که چگونه معیارهای تشخیصی موثر هنگامی که بر اساس ارزش‌های حرفه‌ای پزشکان اندازه‌گیری شوند، می‌توانند وقوع اشتباهات پزشکی را کاهش دهند. موفقیت نظام پزشکی جیسینجر نشان‌دهنده پیشرفت‌های چشمگیری است که با استفاده از اندازه‌گیری کامپیوتری همراه با فرهنگ نهادی مبتنی بر همکاری ادغام شده است؛ جایی که تعیین و تنظیم معیارهای اندازه‌گیری و ارزیابی عملکرد توسط تیم‌های پزشکان و مدیران انجام می‌شود. در هر دو مورد، از معیارها به‌گونه‌ای استفاده شده است که برای حرفه‌ای‌گرایی و انگیزه‌های ذاتی جذاب باشد. اما همان‌طور که دیدیم در جاهای دیگر نظام پزشکی، استفاده از پاداش برای عملکرد اندازه‌گیری شده گاهی بی‌ثمر بوده و منجر به خروجی‌های برعکس شده است.

بازتاب‌ها درباره بهترین استفاده از معیارهای عملکرد توسط ارتش آمریکا در مبارزات عملیات ضد شورش نشان داد در حالی که معیارهای استاندارد اغلب فریبنده هستند، معیارهایی که با موردهای متناسب به‌ویژه توسط افراد حرفه‌ای با تجربه محلی توسعه یافته‌اند، می‌توانند واقعا آموزنده باشند. چالش در چنین مواردی، کنار گذاشتن الگوهای

1. Peter Pronovost's Keystone

جهانی و کشف آن چیزی است که ارزش شمارش دارد و اعداد و ارقامی که در زمینه محلی خود معنادار است. همان‌طور که بارها و بارها دیده‌ایم اندازه‌گیری جایگزین قضاوت نیست؛ اندازه‌گیری نیازمند قضاوت است: قضاوت درباره اینکه آیا اندازه‌گیری شود، چه چیزی اندازه‌گیری شود، چگونه پاداش‌ها و تنبیهات به نتایج اندازه‌گیری می‌شود ارزیابی شود، چگونه پاداش‌ها و تنبیهات به نتایج پیوست شود و اینکه چه کسی اندازه‌گیری‌ها را در دسترس قرار دهد. اگر خود را در موقعیت تنظیم سیاست می‌بینید، در اینجا پرسش‌هایی مطرح است که باید بررسیید و عواملی که باید در نظر داشته باشید؛ در رابطه با اینکه آیا باید از اندازه‌گیری عملکرد استفاده کرد و اگر چنین است، چگونه چنین چیزی باید انجام شود. این پرسش‌ها فهرستی از اندازه‌گیری عملکرد موفقیت را ایجاد کرده‌اند. با توجه به آنچه درباره خطرات ثبات معیارها گفتیم باید این نکته را در نظر گرفت که ممکن است بهترین استفاده از معیارها این باشد که اصلاً از آنها استفاده نشود.

## فهرست

۱. فکر می‌کنید چه نوع اطلاعاتی را می‌توانید اندازه‌گیری کنید؟ هر چه چیزی که اندازه‌گیری می‌شود بیشتر شبیه اشیای بی‌جان باشد، امکان قابل اندازه‌گیری بودن آن نیز بیشتر و به همین دلیل اندازه‌گیری در علوم طبیعی و مهندسی ضروری است. وقتی اشیایی که قرار است اندازه‌گیری شوند تحت تاثیر فرآیند اندازه‌گیری قرار می‌گیرند، اندازه‌گیری کمتر قابل اعتماد می‌شود. هنگامی که فعالیت انسان اندازه‌گیری می‌شود، اندازه‌گیری بسیار کمتر قابل اعتماد می‌شود، زیرا افراد نسبت به خویشتن آگاهی دارند و می‌توانند به فرآیندی که قرار است اندازه‌گیری شود، واکنش نشان دهند. اگر پاداش‌ها و تنبیهات نیز در این فرآیند درگیر باشد، احتمالاً به شیوه‌ای واکنش نشان می‌دهند که اعتبار اندازه‌گیری زیر سوال رود. برعکس، هر چقدر آنها موافق اهداف پاداش‌ها و جوایز

باشند، احتمال بیشتری دارد که به گونه‌ای واکنش نشان دهند که اعتبار اندازه‌گیری بیشتر شود.

۲. چگونه اطلاعات سودمند است؟ همیشه با یادآوری این واقعیت آغاز کنید که بعضی فعالیت‌های قابل اندازه‌گیری، ارزش اندازه‌گیری ندارند. در واقع، سهولت اندازه‌گیری شاید رابطه معکوسی با آنچه که اندازه‌گیری می‌شود داشته باشد. به عبارت دیگر، از خود پرسید آیا چیزی که در حال اندازه‌گیری آن هستید واقعا نماینده چیزیست که می‌خواهید درباره‌اش بدانید؟ اگر اطلاعات، سودمند یا نماینده راستین آنچه که می‌خواهید درباره‌اش بدانید نیست، بهتر است سراغ اندازه‌گیری نروید.

۳. چگونه معیارهای بیشتر سودمند است؟ به یاد داشته باشید که عملکرد اندازه‌گیری شده اگر سودمند باشد در شناخت عوامل بیرونی به‌ویژه عملکردهای ضعیف یا سوءرفتارهای حقیقی موثرتر است. هرچند احتمالاً برای تمایز میان افرادی که در وسط یا بالای نردبان عملکرد هستند، کمتر مفید خواهد بود. به علاوه، هر چقدر شما بیشتر اندازه‌گیری کنید احتمال اینکه هزینه‌های جانبی اندازه‌گیری بیشتر از مزایا شود نیز بیشتر است. در واقع گزاره «معیارها مفید است» بدین معنا نیست که معیارهای بیشتر، کمک‌کننده‌تر باشد.

۴. اتکا نکردن به اندازه‌گیری استاندارد چه هزینه‌هایی دارد؟ آیا به جز قضاوت و تجربه مشتریان، بیماران یا والدین دانش آموزان، منابع اطلاعاتی دیگری درباره عملکرد وجود دارد؟ مثلاً در مورد مدرسه، میزان درخواست والدین برای فرزندان‌شان از یک معلم خاص احتمالاً شاخص مفیدی است که معلم کار خود را درست انجام می‌دهد، حال چه نتایج در آزمون‌های استاندارد نشان داده شود چه نشود. در مورد خیریه‌ها، چه‌بسا سودمندترین حالت این باشد که به خیرین اجازه قضاوت درباره نتایج را داد.

۵. اندازه‌گیری با چه هدفی انجام می‌شود یا به عبارت دیگر

اطلاعات برای چه کسانی شفاف‌سازی خواهد شد؟ در اینجا یک تمایز مهم وجود دارد: میان داده‌هایی که با اهداف ارزیابی داخلی عملکرد توسط خود متخصصان استفاده می‌شود و داده‌هایی که توسط بخش‌های خارجی برای پاداش و تنبیه مورد استفاده قرار می‌گیرد. مثلاً، آیا از داده‌های جرایم برای کشف اینکه نیروهای پلیس باید در کجا جوخه‌های ماشینی بیشتری را مستقر کنند یا تصمیم درباره اینکه فرمانده حوزه ارتقا بگیرد، استفاده می‌شود؟ یا اینکه یک تیم جراحی از داده‌ها استفاده می‌کند برای کشف اینکه کدام فرآیندها بهترین عمل را دارند یا مدیران داده‌های مشابهی را مورد استفاده قرار می‌دهند تا تصمیم بگیرند آیا به بیمارستان بر اساس امتیازاتش پاداش دهند یا آن را تنبیه کنند؟ ابزارهای اندازه‌گیری مانند آزمون‌ها بسیار ارزشمند هستند، اما آنها برای تحلیل‌های داخلی متخصصان بیشترین فایده را دارند تا برای ارزیابی خارجی مخاطبان عمومی که ممکن است نتوانند محدودیت‌ها را درک کنند. از این اندازه‌گیری می‌توان برای آگاه‌سازی پزشکان از عملکرد خود نسبت به هم‌تایان‌شان استفاده کرد و کسانی که سرآمد هستند می‌توانند به افرادی که عقب مانده‌اند کمک کنند. این رویه تا آنجا می‌تواند بسط یابد که آنها درگیر تعیین اشتغال و پرداخت و به تدریج وارد روش‌های بازی‌های آماری یا تقلب‌های آشکار می‌شوند.

به یاد داشته باشید، همان‌طور که دیدیم معیارهای عملکردی که پاداش و مجازات را مرتبط می‌سازند چه‌بسا واقعاً به تقویت انگیزه‌ها و محرک‌ها کمک کنند، به‌ویژه هنگامی که اهدافی که به آنها پاداش داده می‌شود با اهداف حرفه‌ای شاغلان مطابقت داشته باشد. [۱] از سوی دیگر، هرچند منظور از طرح پاداش و مجازات ارائه رفتار روشن از سوی شاغلان است، اما اغلب اوقات آسیب‌زا و بی‌فایده است زیرا احتمال بیشتری وجود دارد تا معیارها به روش‌های مختلفی دستکاری شود و اگر شاغلان تمایل زیادی

به پاداش‌های خارجی داشته باشند، شاید به قیمت غفلت از سایر جنبه‌هایی که به همان اندازه اهمیت دارد، تمرکز خود را روی فعالیت‌هایی بگذارند که اندازه‌گیری و پاداش داده می‌شود. به این دلایل، معیارهای «کم‌خطر»، اغلب موثرتر از معیارهای پرخطر است. پرداخت مبتنی بر عملکرد را به خاطر بیاورید که در آنها افراد با انگیزه‌های بیرونی بیشتر تشویق می‌شوند تا انگیزه‌های درونی؛ این بدان معناست که افراد به کسب درآمد بیشتر از سایر مزایای بالقوه کاری، اجتماعی و فکری اهمیت می‌دهند. این موضوع ممکن است به این دلیل باشد که در برخی حوزه‌ها مانند مالیه افراد موفقیت حرفه‌ای خود را تقریباً با پول کسب‌شده می‌سنجند (همان‌طور که پیشتر گفتیم، این امر مانع نمی‌شود که افراد از درآمد خود برای طیف گسترده‌ای از اهداف از جمله خودخواهانه استفاده کنند). هنگامی که شغل، جذابیت‌های چندانی ندارد - زمانی که به طور پیوسته تکرار می‌شود و فضای اندکی برای انتخاب باقی می‌ماند، مثل تعویض شیشه جلو اتومبیل یا تهیه همبرگر - پرداخت مبتنی بر عملکرد اندازه‌گیری شده شاید اثربخش باشد.

۶. هزینه‌های به دست آوردن معیارها چیست؟ به دست آوردن اطلاعات هرگز رایگان نیست و اغلب به روش‌هایی که به‌ندرت برای خواستاران آن اتفاق می‌افتد بسیار پرهزینه است. جمع‌آوری داده‌ها، پردازش و تجزیه و تحلیل آنها؛ همه اینها زمان‌بر و هزینه‌های آنها هزینه‌های فرصت زمانی است که به آنها اختصاص می‌یابد. به عبارت دیگر، هر لحظه‌ای که شما یا همکاران یا کارمندان شما به تولید معیارها اختصاص دهید، زمان اختصاص داده‌شده به فعالیت‌های اندازه‌گیری نیست. البته اگر شما تحلیلگر داده هستید، تولید معیارها فعالیت اصلی شماست اما برای دیگر افراد این یک حواس‌پرتی به شمار می‌رود. بنابراین حتی اگر عملکردها ارزش اندازه‌گیری داشته باشند ممکن است ارزش آنها کمتر از هزینه‌های آنها باشد. همچنین به یاد داشته باشید که هزینه‌های زمان و

تلاش انسان نیز غیرقابل محاسبه و این دلیل دیگری برای محتاط بودن است.

۷. برسید چرا افرادی که در رأس سازمانند به دنبال معیارهای عملکرد هستند. همان‌طور که اشاره کردیم، تقاضا برای اندازه‌گیری عملکرد گاهی ناشی از ناآگاهی مدیران نسبت به نهادی است که برای مدیریت آن استخدام شده‌اند و این ناآگاهی نتیجه منصوب شدن در سازمانی است که تجربه اندکی در آن دارند. از آنجا که تجربه و دانش محلی مهم است، به سوی استخدام از درون بروید. حتی اگر فردی باهوش‌تر و موفق‌تر در جای دیگری یافتید، نداشتن دانش لازم درباره شرکت، دانشگاه، سازمان دولتی یا دیگر سازمان‌ها چه‌بسا از مزایای استخدام او بیشتر نباشد.

۸. چگونه و توسط چه کسی (چه کسانی) معیارهای عملکرد ایجاد شده و گسترش یافته‌اند؟ هنگامی که معیارهای پاسخگویی با فرمول‌های استانداردسازی شده و توسط افراد دور از تعامل با فعالیت‌های در حال انجام و از بالا تحمیل می‌شود، احتمال اینکه موثر واقع شود بسیار کمتر است. این به معنای درخواست از صاحبان دانش ضمنی مبتنی بر تجربه مستقیم است تا در چگونگی ایجاد و گسترش استانداردهای عملکرد مناسب دخیل شوند. [۲] سعی کنید گروهی را به نمایندگی از همه افرادی که در خروجی‌ها مشارکت دارند انتخاب کنید. [۳] در بهترین حالت، آنها باید همواره به عنوان بخشی از فرآیند ارزیابی داده‌های اندازه‌گیری شده حضور داشته باشند.

به یاد داشته باشید که سیستم عملکرد اندازه‌گیری شده تا اندازه‌ای به کار گرفته می‌شود که افراد به ارزش آن باور دارند. از همین رو، ما در این فصل تصمیم گرفتیم که به بررسی چشم‌انداز افرادی پردازیم که در موقعیت تصمیم‌گیری در این رابطه قرار دارند که چرا و چگونه معیارهای نهادی کار می‌کنند اما اگر چنین موقعیتی نداشته باشید و در سلسله‌مراتب سازمانی در رده‌های پایین باشید و یک مدیر میانی یا رئیس یک بخش دانشگاهی از شما انتظار

اجرای معیارها را داشته باشد، چه می‌شود؟ آن زمان چه می‌کنید؟ آن وقت با یک انتخاب روبه‌رو می‌شوید. اگر به اهدافی که اطلاعات برای آنها جمع‌آوری شده است باور دارید، پس چالش شما تهیه و ارائه داده‌های دقیق با کاراترین روش ممکن است؛ اطلاعاتی که حداقل زمان شما و افراد تحت مدیریت شما را می‌طلبد. برعکس، اگر درباره اهداف تردید داشته باشید و بر این باور باشید که آنها بی‌فایده‌اند، شاید برای متقاعد کردن مافوق خود تلاش کنید (شاید با دادن یک نسخه از این کتاب). در صورت دست نیافتن به موفقیت، وظیفه شما تهیه و ارائه داده‌ها به روشی است که کمترین زمان را صرف کند، حداقل استانداردهای قابل پذیرش را داشته باشد و به واحد شما آسیب نرساند.

اگر شما در رأس یا نزدیک به رأس سازمان هستید و تصمیم‌گیری درباره معیارها با شماست، پاراگراف قبل را دوباره بخوانید و به خاطر داشته باشید که روش‌های متفاوتی برای واکنش نشان دادن افراد زیردست شما وجود دارد. معیارها زمانی که درگیر اهداف و اعتبار شوند، بهترین عملکرد خود را دارند. [۴]

۹. به یاد داشته باشید که حتی بهترین معیارها نیز موجب فساد یا انحراف اهداف هستند. از آنجا که افراد عاملان به حداکثر رساندن منافع خود هستند، اشکالات گریزناپذیری در همه طرح‌های پاداش مبتنی بر اندازه‌گیری وجود دارد. اگر در حال حاضر این‌طور باشد که به پزشکان بر اساس فرآیند عملکردشان پاداش داده می‌شود، این انگیزه را برای آنها ایجاد می‌کند که فرآیندهایی را انجام دهند که هزینه‌های بالا و مزایای اندکی دارد. اما اگر به پزشکان بر اساس تعداد بیمارانی که ویزیت می‌کند پاداش داده شود، در این صورت آنها انگیزه‌شان برای پذیرش بیماران بیشتر افزایش می‌یابد و فرآیندهای زمان‌گیر اما بی‌هوده را کاهش می‌دهند. اگر به آنها بر اساس خروجی‌های بیماران‌شان پاداش دهید، در این صورت احتمال کار سطحی و پرهیز از پذیرش و معالجه بیماران مشکل‌دارتر، بیشتر می‌شود. [۵]

این بدان معنا نیست که معیارهای عملکرد به این دلیل که برخی خروجی‌های منفی دارد باید به طور کلی کنار گذاشته شود. با وجود مشکلات قابل پیش‌بینی، معیارها هنوز ارزش استفاده دارد: موضوع، سبک سنگین کردن و البته قضاوت است.

۱۰. به یاد داشته باشید که گاهی شناخت محدودیت‌های ممکن، سرآغاز دانایی است. همه مشکلات قابل حل نیست و تعداد بسیار کمتری از آنها با معیارها حل می‌شود. این تصور درستی نیست که همه چیز با اندازه‌گیری بهتر می‌شود یا در حقیقت هر چیزی را که بتوان اندازه‌گیری کرد، بهبود خواهد یافت. شفاف‌تر شدن یک مشکل لزوماً گامی به سوی حل آن مشکل نیست. شفافیت صرفاً ممکن است یک وضعیت آشفته را تا اندازه‌ای آرام کند اما راه حلی ارائه نمی‌دهد.

در پایان باید گفت راه حل ساده و جادویی برای مشکلات وجود ندارد، هیچ جایگزینی برای شناخت یک موضوع یا سازمان نیست و این موضوع بخشی از تجربه و مهارت غیرقابل اندازه‌گیری است. بسیاری از مسائل و موضوعات مهم بیش از حد مورد تفسیر و قضاوت قرار گرفته‌اند که بتوان آنها را با معیارهای استاندارد حل کرد. سرانجام، مساله، معیارها در برابر قضاوت نیست، بلکه معیارها داوری آگاه‌کننده‌اند که در بر دارنده دانشی درباره میزان وزنی است که این داوری‌ها به معیارها می‌دهد، و همین‌طور حاوی شناخت ویژگی‌های تحریف‌کننده معیارها و درک مواردی است که نمی‌توان آنها را اندازه‌گیری کرد. در دهه‌های اخیر، بسیاری از سیاست‌مداران، صاحبان کسب‌وکار، سیاست‌گذاران و مقامات دانشگاهی از این موضوع چشم‌پوشی کرده‌اند.



## یادداشت‌ها

## پیشگفتار

1. Gwyn Bevan and Christopher Hood, "What's Measured Is What Matters: Targets and Gaming in the English Public Health System," *Public Administration* 84, no. 3 (2006), pp. 517–53.
2. Paula Chatterjee and Karen E. Joynt, "Do Cardiology Quality Measures Actually Improve Patient Outcomes?" *Journal of the American Heart Association* (February 2014). The same problem was noted some years earlier by Richard Rothstein, "The Influence of Scholarship and Experience in Other Fields on Teacher Compensation Reform," pp. 87–110 in Matthew G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education* (Washington, D.C., 2009), p. 96; an expanded version was published as *Holding Accountability to Account: How Scholarship and Experience in Other Fields Inform Exploration of Performance Incentives in Education*, National Center on Performance Incentives, Working Paper 2008–04, February 2008.
3. Bevan and Hood, "What's Measured Is What Matters."
4. An exception is Richard Rothstein, *Holding Accountability to Account*. Also valuable is Adrian Perry, "Performance Indicators: 'Measure for Measure' or 'A Comedy of Errors'?" in Caroline Mager, Peter Robinson, et al. (eds.), *The New Learning Market* (London, 2000).
5. Laura Landro, "The Secret to Fighting Infections: Dr. Peter Pronovost Says It Isn't That Hard. If Only Hospitals Would Do It," *Wall Street Journal*, March 28, 2011, and Atul Gawande, *The Checklist Manifesto* (New York, 2009).
6. Michael Lewis, *Moneyball: The Art of Winning an Unfair Game* (New York, 2003).
7. Chris Lorenz, "If You're So Smart, Why Are You under Surveillance? Universities, Neoliberalism, and New Public Management," *Critical Inquiry* (Spring 2012), pp. 599–29, esp. pp. 610–11.

8. Jonathan Haidt, *The Righteous Mind* (New York, 2012), p. 34 and passim.
9. On the Spellings Commission report, see Fredrik deBoer, *Standardized Assessments of College Learning Past and Future* (Washington, D.C.: New American Foundation, March 2016).
10. Jerry Z. Muller, *The Mind and the Market: Capitalism in Modern European Thought* (New York, 2002) and my Teaching Company lecture course, "Thinking about Capitalism." See also Robert K. Merton, "The Unanticipated Consequences of Purposive Social Action," *American Sociological Review* 1 (December 1936), pp. 894–904; and Merton, "Unanticipated Consequences and Kindred Sociological Ideas: A Personal Gloss," in Carlo Mongardini and Simonetta Tabboni (eds.), *Robert K. Merton and Contemporary Sociology* (New Brunswick, N.J., 1998), pp. 295–318; Robert K. Merton and Elinor Barber, *The Travels and Adventures of Serendipity* (Princeton, 2004).

۱۱. همان گونه که الفی کوهن یادآور شده است: «در همان اثنا که روانشناسان اجتماعی آغاز به شناخت چگونگی محرک‌های بیرونی ضد تولید کردند، این پیام از مطالب منتشرشده در حوزه مدیریت حذف شد.»

Kohn, *Punished by Rewards* (New York, 1999), p. 121.

## فصل ۱

1. A term used by Bruce G. Charlton, "Audit, Accountability, Quality and All That: The Growth of Managerial Technologies in UK Universities," in Stephen Prickett and Patricia Erskine-Hill (ed.), *Education! Education! Education! Managerial Ethics and the Law of Unintended Consequences* (Thorverton, England, 2002).
2. Fabrizio Ferraro, Jeffrey Pfeffer, and Robert L. Sutton, "Economics Language and Assumptions: How Theories Can Become Self-Fulfilling," *Academy of Management Review* 30, no. 1 (2005), pp. 8–24.
3. Tom Peters, "What Gets Measured Gets Done," (1986), <http://tompeters.com/columns/what-gets-measured-gets-done/>.
۴. فرمول‌بندی اخیر را مدیون پروفیسور پاول کولیر هستم.
5. Charlton, "Audit, Accountability, Quality and All That," pp. 18–22.
6. Useful attempts to summarize these negative consequences include Colin Talbot, "Performance Management," pp. 491–517 in Ewan Ferlie, Laurence E. Lynn, Jr., and Christopher Pollitt (eds.), *The Oxford Handbook of Public Management* (New York, 2005), pp. 502–4; and Michael Power, "The Theory of the Audit Explosion," pp. 326–44, in the same volume, see esp. p. 335.
7. William Bruce Cameron, *Informal Sociology: A Casual Introduction to Sociological Thinking* (New York, 1963).
8. Bevan and Hood, "What's Measured Is What Matters."
9. Quoted in Diane Ravitch, *The Death and Life of the Great American School System* (New York, 2010), p. 160. See also Chris Shore, "Audit Culture and Illiberal Governance: Universities and the Culture of Accountability," *Anthropological Theory* 8, no. 3

(2008), pp. 278–99; Mary Strathern (ed.), *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy* (London, 2000).

10. Alison Wolf, *Does Education Matter? Myths about Education and Economic Growth* (London, 2002), p. 246. C.A.E. Goodhart, “Problems of Monetary Management: The UK Experience” (1975), pp. 91–121 in Goodhart, *Monetary Theory and Practice* (London, 1984).

## فصل ۲

1. Kurt C. Strange and Robert L. Ferrer, “The Paradox of Family Care,” *Annals of Family Medicine* 7, no. 4 (July/August 2009), pp. 293–99, esp. p.295.
2. Sally Engle Merry, *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking* (Chicago, 2016), pp. 1–33.
3. *Ibid.*, p. 1.
4. Kevin E. Davis, Benedict Kingsbury, and Sally Engle Merry, “Introduction: Global Governance by Indicators,” in Kevin Davis, Angelina Fisher, Benedict Kingsbury, and Sally Engle Merry (eds.), *Governance by Indicators: Global Power through Quantification and Rankings* (New York, 2012), pp. 9, 18.

## فصل ۳

1. Quoted in Matthew Arnold, “The Twice-Revised Code” (1862), in R. H. Super (ed.), *The Complete Prose Works of Matthew Arnold* (Ann Arbor, Mich., 1960–77), vol. 2, pp. 214–15.
2. Park Honan, *Matthew Arnold: A Life* (Cambridge, Mass., 1983), pp. 318–19; R. H. Super’s notes to Arnold, “The Twice-Revised Code,” in *Complete Prose Works*, vol. 2, p. 349.
3. Arnold, “The Twice-Revised Code,” pp. 223–24.
4. *Ibid.*, p. 226.
5. *Ibid.*, p. 243.
6. Fred G. Walcott, *The Origins of Culture and Anarchy: Matthew Arnold and Popular Education in England* (Toronto, 1970), pp. 7–8.
7. Arnold, “Special Report on Certain Points Connected with Elementary Education in Germany, Switzerland, and France” (1886), in *Complete Prose Works*, vol. 11, pp. 1, 28.
8. Simon Patten, “An Economic Measure of School Efficiency,” *Educational Review* 41 (May 1911), pp. 467–69, quoted in Raymond E. Callahan, *Education and the Cult of Efficiency* (Chicago, 1962), p. 48.
9. Frederick W. Taylor, *The Principles of Scientific Management* (New York, 1911). On Taylor and his influence on education reform advocates see Callahan, *Education and the Cult of Efficiency*, chap. 2.
10. Alfred D. Chandler, Jr., *The Visible Hand: The Managerial Revolution in American Business* (Cambridge, Mass., 1977), pp. 275–76.
11. Taylor, quoted in James C. Scott, *Seeing Like a State: How Certain Schemes to Improve*

- the Human Condition Have Failed* (New Haven, 1998), p. 336.
12. Frederick W. Taylor, *Principles of Scientific Management*, cited by David Montgomery, *The Fall of the House of Labor* (New Haven, 1989), p. 229.
  13. Ellwood P. Cubberley, *Public School Administration* (Boston, 1916), on which see Callahan, *Education and the Cult of Efficiency*, pp. 95–99.
  14. Dana Goldstein, *The Teacher Wars: A History of America's Most Embattled Profession* (New York, 2014), pp. 86–87.
  15. For the term “student growth,” see Chad Aldeman, “The Teacher Evaluation Revamp, in Hindsight,” *EducationNext* 17, no. 2 (Spring 2017), <http://educationnext.org/the-teacher-evaluation-revamp-in-hindsight-obama-administration-reform/>.
  16. Richard Sennett, *The Corrosion of Character: The Personal Consequences of Work in the New Capitalism* (New York, 1998), p. 42.
  17. Rakesh Khurana, *From Higher Aims to Hired Hands: The Social Transformation of American Business Schools and the Unfulfilled Promise of Management as a Profession* (Princeton, 2007), p. 295.
  18. Richard R. Locke and J.-C. Spender, *Confronting Managerialism: How the Business Elite and Their Schools Threw Our Lives out of Balance* (London, 2011), p. xiii.
  19. Adrian Wooldridge, *Masters of Management* (New York, 2011), p. 3.
  20. Bob Lutz, *Car Guys vs. Bean Counters: The Battle for the Soul of American Business* (New York, 2013).
  21. Christopher Pollitt, “Towards a New World: Some Inconvenient Truths for Anglosphere Public Administration,” *International Review of Administrative Sciences* 81, no. 1 (2015), pp. 3–17, esp. pp. 4–5; John Quiggin, “Bad Company: Correspondence,” *Quarterly Essay.com*, <https://www.quarterlyessay.com.au/correspondence/1203>; similarly Henry Mintzberg, “Managing Government, Governing Management,” *Harvard Business Review*, May–June 1996, pp. 75–83.
  22. David Halberstam, *The Best and the Brightest* (New York, 1972), pp. 213–65.
  23. Kenneth Cukier and Viktor Mayer-Schönberger, “The Dictatorship of Data,” *MIT Technology Review*, May 31, 2013.
  24. Edward N. Luttwak, *The Pentagon and the Art of War* (New York, 1984), p. 269.
  25. Luttwak, *The Pentagon and the Art of War*, pp. 30–31. On the misuse of the metric of body counts, see the memoirs and scholarly literature analyzed in Ben Connable, *Embracing the Fog of War: Assessment and Metrics in Counterinsurgency* (Rand Corporation, 2012), pp. 106ff.
  26. Luttwak, *The Pentagon and the Art of War*, pp. 138–43.
  27. *Ibid.*, p. 152.
  28. Matthew Stewart, *The Management Myth: Why the Experts Keep Getting It Wrong* (New York, 2009), p. 31.
  29. Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton, 1995), p. ix.

## فصل ۴

1. Ralf Dahrendorf, *The Modern Social Conflict: An Essay on the Politics of Liberty* (Berkeley, 1988), p. 53.
2. Porter, *Trust in Numbers*, p. ix.
3. Stefan Collini, "Against Prodspeak," in Collini, *English Pasts: Essays in History and Culture* (Oxford, 1999), p. 239.
4. Philip K. Howard, *The Rule of Nobody: Saving America from Dead Laws and Broken Government* (New York, 2014), p. 44.
5. Philip K. Howard, *The Death of Common Sense: How Law Is Suffocating America* (New York, 1994), pp. 12, 27.
6. Howard, *The Rule of Nobody*, p. 54.
7. Lawrence M. Freedman, "The Litigation Revolution," in Michael Grossman and Christopher Tomlins (eds.), *The Cambridge History of Law in America: Vol. III The Twentieth Century and After* (Cambridge, 2008), p. 176.
8. *Ibid.*, p. 187.
9. *Ibid.*, p. 188–89.
10. Mark Schlesinger, "On Values and Democratic Policy Making: The Deceptively Fragile Consensus around Market-Oriented Medical Care," *Journal of Health Politics, Policy and Law* 27, no. 6 (December 2002), pp. 889–925; and Mark Schlesinger, "Choice Cuts: Parsing Policymakers' Pursuit of Patient Empowerment from an Individual Perspective," *Health, Economics, Policy and the Law* 5 (2010), pp. 365–87.
11. James Heilbrun, "Baumol's Cost Disease," in Ruth Towse (ed.), *A Handbook of Cultural Economics*, 2nd ed. (Cheltenham, England, 2011); and William G. Bowen, "Costs and Productivity in Higher Education," The Tanner Lectures, Stanford University, October 2012, pp. 3–4.
12. Bowen, "Costs and Productivity in Higher Education," p. 5, citing Teresa A. Sullivan et al. (eds.), *Improving Measurement of Productivity in Higher Education* (Washington, D.C., 2012).
13. Yves Morieux and Peter Tollman, *Six Simple Rules: How to Manage Complexity Without Getting Complicated* (Boston, 2014), p. 6.
14. Rakesh Khurana, *Searching for a Corporate Savior: The Irrational Quest for Charismatic CEOs* (Princeton, 2002), esp. chap. 3. The phenomenon is by no means confined to the corporate sector.
15. Steven Levy, "A Spreadsheet Way of Knowledge," *Harper's*, November 1984, now online at <https://medium.com/backchannel/a-spreadsheet-way-of-knowledge-8de60af7146e>.
16. Seth Klarman, *A Margin of Safety: Risk-Averse Value Investing for the Thoughtful Investor* (New York, 1991).

## فصل ۵

1. Michael Jensen and William H. Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics* 3, no. 4 (1976), pp. 305–60; Bengt Holmström and Paul Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, & Organization* [Special Issue: Papers from the Conference on the New Science of Organization, January 1991] 7 (1991), pp. 24–52; Charles Wheelan, *Naked Economics*, rev. ed. (New York, 2010), pp. 39–43.
2. Khurana, *From Higher Aims to Hired Hands*, pp. 317–26. Similarly, Richard Münch, *Globale Eliten, lokale Autoritäten* (Frankfurt, 2009), p. 75. Also instructive is Ferraro, Pfeffer, Sutton, "Economics Language and Assumptions."
3. Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton, 1995), p. ix.
4. Talbot, "Performance Management," p. 497.
5. David Chinitz and Victor G. Rodwin, "What Passes and Fails as Health Policy and Management," *Journal of Health Politics, Policy, and Law* 39, no. 5 (October 2014), pp. 1113–26, esp. pp. 1114–17.
6. Mintzberg, "Managing Government, Governing Management," pp. 75–83; and Holmström and Milgrom, "Multitask Principal-Agent Analyses."
7. Hal K. Rainey and Young Han Chun, "Public and Private Management Compared," in Ewan Ferlie, Laurence E. Lynn, Jr., and Christopher Pollitt (eds.), *The Oxford Handbook of Public Management* (New York, 2005), pp. 72–102, 85; and James Q. Wilson, *Bureaucracy: What Government Agencies Do and Why They Do It* (New York, 2000), pp. 156–57.
8. Roland Bénabou and Jean Tirole, "Intrinsic and Extrinsic Motivation," *Review of Economic Studies* no. 70 (2003), pp. 489–520. A pioneering work of intrinsic motivation theory was Edward L. Deci, *Intrinsic Motivation* (New York, 1975). Other studies by psychologists include Thane S. Pittman, Jolee Emery, and Ann K. Boggiano, "Intrinsic and Extrinsic Motivational Orientations: Reward-Induced Changes in Preference for Complexity," *Journal of Personality and Social Psychology* 42, no. 5 (1982), pp. 789–97; and T. S. Pittman, A. K. Boggiano, and D. N. Ruble, "Intrinsic and Extrinsic Motivational Orientations: Limiting Conditions on the Undermining and Enhancing Effects of Reward on Intrinsic Motivation," in J. Levine and M. Wang (eds.), *Teacher and Student Perceptions: Implications for Learning* (Hillsdale, N.J., 1983). An important figure in the transition of the theory from psychology to economics is Bruno S. Frey, whose works include *Not Just for the Money: An Economic Theory of Human Motivation* (Cheltenham, England, 1997).

برای مرور ادبیات مربوطه با محوریت اقتصاد رفتاری که نتیجه می‌گیرد «اقتصاد رفتاری به وضوح نشان می‌دهد که کاربرد جهانی پرداخت مبتنی بر عملکرد که امروزه در حال اجراست مستلزم حقایق علمی نیست.» برای مطالعه بیشتر در این زمینه به منابع زیر مراجعه کنید:

Antoinette Weibel, Meike Wiemann, and Margit Osterloh, "A Behavioral Economics Perspective on the Overjustification Effect: Crowding-In and Crowding-Out

- of Intrinsic Motivation,” in Marylène Gagné (ed.), *The Oxford Handbook of Work Engagement, Motivation, and Self-Determination Theory* (New York, 2014).
9. Pittman, Boggiano, and Ruble, “Intrinsic and Extrinsic Motivational Orientations.”
  10. Bénabou and Tirole, “Intrinsic and Extrinsic Motivation,” p. 504.
  11. Bruno S. Frey and Margit Osterloh, “Motivate People with Prizes,” *Nature* 465, no. 17 (June 2010), p. 871.
  12. George Akerlof, “Labor Contracts as a Partial Gift Exchange,” *Quarterly Journal of Economics* 97, no. 4 (1982), 543–69.
  13. Bruno S. Frey and Reto Jegen, “Motivation Crowding Theory,” *Journal of Economic Surveys* 15, no. 5 (2001), pp. 589–611; and Robert Gibbons, “Incentives in Organizations,” *Journal of Economic Perspectives* 12, no. 4 (Fall 1998), pp. 115–32, esp. p. 129.
  14. Gibbons, “Incentives in Organizations.”
  15. Talbot, “Performance Management,” pp. 491–517; Adrian Wooldridge, *Masters of Management* (New York, 2011), pp. 318–19; Pollitt, “Towards a New World”; Christopher Hood, “The ‘New Public Management in the 1980s: Variations on a Theme,” *Accounting, Organization, and Society* 20, nos. 2/3 (1995), pp. 93–109; Christopher Hood and Guy Peters, “The Middle Aging of New Public Management: Into the Age of Paradox?” *Journal of Public Administration Research and Theory* 14, no. 3 (2004), pp. 267–82. On the background and early history of NPM in the United Kingdom and the United States, see Christopher Pollitt, *Managerialism and the Public Services*, 2nd ed. (Oxford, 1993).

## فصل ۶

1. Harry Braverman, *Labor and Monopoly Capital* (New York, 1974).
2. Michael Oakeshott, “Rationalism in Politics” (1947) in Oakeshott, *Rationalism in Politics and Other Essays* (Indianapolis, 1991).
3. Friedrich Hayek, “The Uses of Knowledge in Society,” “The Meaning of Competition,” and “‘Free’ Enterprise and Competitive Order,” all in Hayek, *Individualism and Economic Order* (Chicago, 1948).
4. Wolf, *Does Education Matter?* p. 246; Lorenz, “If You’re So Smart”; Bevan and Hood, “What’s Measured Is What Matters.” For an extended analysis of the ways in which the British higher education regime replicates features of the Soviet system, see Aviezer Tucker, “Bully U: Central Planning and Higher Education,” *Independent Review* 17, no. 1 (Summer 2012), pp. 99–119.
5. Alfie Kohn, *Punished by Rewards* (New York, 1999), pp. 62ff; and Teresa Amabile, “How to Kill Creativity,” *Harvard Business Review* (September–October 1998).
6. Scott, *Seeing Like a State*, p. 313.
7. Isaiah Berlin, “Political Judgment,” in Berlin, *The Sense of Reality: Studies in Ideas and Their History*, ed. Henry Hardy (New York, 1996), pp. 53, 50.

8. Elie Kedourie *Diamonds into Glass: The Government and the Universities* (London, 1988), reprinted in Elie Kedourie, "The British Universities under Duress," *Minerva* 31, no. 1 (March, 1993), pp. 56–105.
9. Elie Kedourie, *Perestroika in the Universities* (London, 1989), pp. x–xi.
10. Kedourie, *Perestroika*, p. 29.
11. Kedourie "The British Universities under Duress," p. 61.
12. Background information on GPRA at <http://www.foreffectivegov.org/node/326>, and Donald P. Moynihan and Stephane Lavertu, "Does Involvement in Performance Management Routines Encourage Performance Information Use? Evaluating GPRA and PART" *Public Administration Review* 72, no. 4 (July/August 2012), pp. 592–602.

## فصل ٧

1. Department of Education, "For Public Feedback: A College Ratings Framework" (December, 2014), <http://www2.ed.gov/documents/collegeaffordability/college-ratings-fact-sheet.pdf>.
2. [https://www.luminafoundation.org/files/publications/stronger\\_nation/2016/A\\_Stronger\\_Nation-2016-National.pdf](https://www.luminafoundation.org/files/publications/stronger_nation/2016/A_Stronger_Nation-2016-National.pdf).
3. Wolf, *Does Education Matter?*
4. Wolf, *Does Education Matter?*; Jason R. Abel, Richard Deitz, and Yaquin Su, "Are Recent College Graduates Finding Good Jobs?" *Federal Reserve Bank of New York: Current Issues in Economics and Finance* 20, no. 1 (2014); Paul Beaudry, David A. Green, Benjamin M. Sand, "The Great Reversal in the Demand for Skill and Cognitive Tasks," NBER Working Paper 18901, March 2013.
5. Katherine Mangan, "High-School Diploma Options Multiply, but May Not Set Up Students for College Success," *Chronicle of Higher Education*, October 19, 2015.
6. Scott Jaschik, "ACT Scores Drop as More Take Test," *Inside Higher Education*, August 24, 2016; and "ACT Scores Down for 2016 U.S. Grad Class Due to Increased Percentage of Students Tested," <http://www.act.org/content/act/en/newsroom/act-scores-down-for-2016-usgrad-class-due-to-increased-percentage-of-students-tested.html>.
7. William G. Bowen and Michael S. McPherson, *Lesson Plan: An Agenda for Change in American Higher Education* (Princeton, 2016), p. 30.
8. Tucker, "Bully U," p. 104.
9. Valen E. Johnson, *Grade Inflation: A Crisis in College Education* (New York, 2003).
10. John Bound, Michael F. Lovenheim, and Sarah Turner, "Increasing Time to Baccalaureate Degree in the United States, NBER Working Paper 15892, April 2010, p. 13; and Sarah E. Turner, "Going to College and Finishing College. Explaining Different Educational Outcomes," in Caroline M. Hoxby (ed.) *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It* (Chicago, 2004), pp. 13–62, <http://www.nber.org/chapters/c10097.passim>.
11. Arnold Kling and John Merrifield, "Goldin and Katz and Education Policy Failings,"

- Econ Journal Watch 6, no. 1 (January 2009), pp. 2–20, esp. p. 14.
12. Wolf, Does Education Matter? chap. 7.
  13. Lorelle L. Espinosa, Jennifer R. Crandall, and Malika Tukibayeva, Rankings, Institutional Behavior, and College and University Choice (Washington, D.C., American Council on Education, 2014), p. 12.
  14. Wolf, Does Education Matter? chap. 7; similarly Daron Acemoglu and David Autor, “What Does Human Capital Do?” *Journal of Economic Literature* 50, no. 2 (2012), pp. 426–63.
  15. Wolf, Does Education Matter? chaps. 2 and 6; Paul Beaudry, David A. Green, and Benjamin M. Sand, “The Great Reversal in the Demand for Skill and Cognitive Tasks,” NBER Working Paper 18901, March 2013.
  16. Stuart Eizenstat and Robert Lerman, “Apprenticeships Could Help U.S. Workers Gain a Competitive Edge” (Washington, D.C., Urban Institute, May 2013); Mark P. Mills, “Are Skilled Trades Doomed to Decline?” Manhattan Institute, New York, 2016, <http://www.manhattaninstitute.org/sites/default/files/IB-MM-1016.pdf>.
  17. Thomas Hale and Gonzalo Viña, “University Challenge: The Race for Money, Students and Status,” *Financial Times*, June 23, 2016; <https://www.oecd.org/unitedkingdom/United%20Kingdom-EAG2014-Country-Note.pdf>. Stefan Collini, in *What Are Universities For?* gives a figure of 45 percent in 2012 for enrollment in higher education.
  18. Stefan Collini, *What Are Universities For?* chap. 2.
  19. Wolf, Does Education Matter? chap. 7.
  20. Shore, “Audit Culture and Illiberal Governance,” pp. 289–90. See also James Wilsdon et al., *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management* (July 2015).
  21. Stephen Prickett, *Introduction to Education! Education! Education!*
  22. Charlton, “Audit, Accountability, Quality and All That,” p. 23.
  23. [http://www2.ed.gov/admins/finaid/accred/accreditation\\_pg6.html](http://www2.ed.gov/admins/finaid/accred/accreditation_pg6.html).
  24. Peter Augustine Lawler, “Truly Higher Education,” *National Affairs* (Spring 2015), pp. 114–30, esp. pp. 120–21.
  25. Charlton, “Audit, Accountability, Quality and All That,” p. 22; and Lorenz, “If You’re So Smart.” p. 609.
  26. Benjamin Ginsberg, *The Fall of the Faculty: The Rise of the All Administrative University* (Baltimore, 2013).
  27. Craig Totterow and James Evans, “Reconciling the Small Effect of Rankings on University Performance with the Transformational Cost of Conformity” in Elizabeth Popp Berman and Catherine Paradeise (eds.), *The University under Pressure*, *Research in the Sociology of Organizations*, vol.
  28. Wendy Nelson Espeland and Michael Sauder, “Rankings and Reactivity: How Public Measures Re-create Social Worlds,” *American Journal of Sociology* 113, no. 1 (July 2007), pp. 1–40, esp. p. 11.

29. Ibid., p. 25.
30. Ibid., p. 26.
31. Ibid., pp. 30–31. For more on how some law schools game the statistics, see Alex Wellen, “The \$8.78 Million Maneuver,” *New York Times*, July 31, 2005. For more on how universities attempt to rise in the rankings, see Wendy Nelson Espeland and Michael Sauder, “The Dynamism of Indicators” in Davis et al. (eds.) *Governance by Indicators*, pp. 103–5.
32. Doug Lederman, “‘Manipulating,’ Er, Influencing ‘U.S. News,’” *Inside Higher Ed*, June 3, 2009.
33. Totterow and Evans, “Reconciling the Small Effect of Rankings on University Performance with the Transformational Cost of Conformity.”
34. On the early history of this, see Collini, *What Are Universities For?* chap. 6, “Bibliometry.”
35. Prickett, Introduction to *Education! Education! Education!* p. 7.
36. Peter Weingart, “Impact of Bibliometrics upon the Science System: Inadvertent Consequences,” *Scientometrics* 62, no. 1 (2005), pp. 117–31, esp. p. 126.
37. Ibid., p. 127; see also Christian Fleck, “Impact Factor Fetishism,” *European Journal of Sociology* 54, no. 2 (2013), pp. 327–56. On the difficulties of comparing research productivity among disciplines, see Dorothea Jansen et al., “Drittmittel als Performanzindikator der wissenschaftlichen Forschung: Zum Einfluss von Rahmenbedingungen auf Forschungsleistung,” *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59, no 1 (2007), pp. 125–49.
38. Weingart, “Impact of Bibliometrics upon the Science System,” and Michael Power, “Research Evaluation in the Audit Society,” in Hildegard Matthies and Dagmar Simon (eds.), *Wissenschaft unter Beobachtung: Effekte und Defekte von Evaluationen* (Wiesbaden, 2008), pp. 15–24.
39. Carl T. Bergstrom, “Use Ranking to Help Search,” *Nature* 465, no. 17 (June 2010), p. 870.
40. Espeland and Sauder, “Rankings and Reactivity,” p. 15. See too Münch, *Globale Eliten, lokale Autoritäten*.
41. Espinosa, Crandall, and Tukibayeva, *Rankings, Institutional Behavior, and College and University Choice*; Douglas Belkin, “Obama Spells Out College-Ranking Framework,” *Wall Street Journal*, December 19, 2014; Jack Stripling, “Obama’s Legacy: An Unlikely Hawk on Higher Ed,” *Chronicle of Higher Education*, September 30, 2016.
42. Jonathan Rothwell, “Understanding the College Scorecard,” paper, Brookings Institution, September 28, 2015, <https://www.brookings.edu/opinions/understanding-the-college-scorecard/>; Beckie Supiano, “Early Evidence Shows College Scorecard Matters, but Only to Some,” *Chronicle of Higher Education*, May 27, 2016.
- 43 Lauren A. Rivera, *Pedigree: How Elite Students Get Elite Jobs* (Princeton, 2015); and Elizabeth A. Armstrong and Laura T. Hamilton, *Paying for the Party: How College Maintains Inequality* (Cambridge, Mass., 2013).
44. Rothwell, “Understanding the College Scorecard.”

45. Jeffrey Steedle, "On the Foundations of Standardized Assessment of College Outcomes and Estimating Value Added," in K. Carey and M. Schneider (eds.), *Accountability in Higher Education* (New York, 2010), p. 8.
46. Nicholas Tampio, "College Ratings and the Idea of the Liberal Arts," *JStOR Daily*, July 8, 2015, <http://daily.jstor.org/college-ratings-idea-liberal-arts/>, and James B. Stewart, "College Rankings Fail to Measure the Influence of the Institution," *New York Times*, October 2, 2015.
47. Robert Grant, "Education, Utility and the Universities" in Prickett and Erskine-Hill (eds.), *Education! Education! Education!* p. 52.
48. Rivera, *Pedigree*, p. 78 and passim.
49. Espinosa, Crandall, and Tukibayeva, *Rankings, Institutional Behavior, and College and University Choice*, p. 9.

## فصل ۸

1. Ravitch, *The Death and Life of the Great American School System*, p. 149.
2. T. S. Dee and B. Jacob, "The Impact of 'No Child Left Behind' on Student Achievement," *Journal of Policy Analysis and Management* 30 (2011), pp. 418–46.
3. Jesse Rhodes, *An Education in Politics: The Origins and Evolution of No Child Left Behind* (Ithaca, N.Y., 2012), p. 88.
4. Quoted in *ibid.*, p. 88.
5. *Ibid.*, p. 153.
6. Goldstein, *Teacher Wars*, p. 188.
7. Diane Ravitch, *Reign of Error* (New York, 2013), p. 51 and charts on pp. 340–42; and Kristin Blagg and Matthew M. Chingos, *Varsity Blues: Are High School Students Being Left Behind?* (Washington, D.C.: Urban Institute, May 2016), pp. 3–5.
8. Dee and Jacob, "The Impact of 'No Child Left Behind' on Student Achievement," pp. 418–46. See also, Ravitch, *The Death and Life of the Great American School System*, pp. 107–8, 159; Goldstein, *Teacher Wars*, p. 187; and American Statistical Association, "ASA Statement on Using Value-Added Models for Educational Assessment, April 8, 2014," [https://www.amstat.org/policy/pdfs/ASA\\_VAM\\_Statement.pdf](https://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf).
9. Goldstein, *Teacher Wars*, p. 226.
10. *Ibid.* For an argument in favor of value-added testing over evaluating schools based on "proficiency" or "college-readiness," see Michael J. Petrilli and Aaron Churchill, "Why States Should Use Student Growth, and Not Proficiency Rates, when Gauging School Effectiveness," Thomas Fordham Institute, October 13, 2016, <https://edexcellence.net/articles/whystates-should-use-student-growth-and-not-proficiency-rates-when-gaugingschool>.
11. Martin R. West, written statement to U.S. Senate Committee on Health, Education, Labor, and Pensions, January 21, 2015, <http://www.help.senate.gov/imo/media/doc/West.pdf>, and David J. Deming et al., "When Does Accountability Work?" *educationnext*.

org (Winter 2016), pp. 71–76. On Florida, David N. Figlio and Lawrence S. Getzler, “Accountability, Ability, and Disability: Gaming the System,” NBER Working Paper No. 9307, October 2002.

12. On Houston and Dallas, see Ravitch, *Death and Life*, p. 155; on Atlanta, see Rachel Aviv, “Wrong Answer: In an Era of High-Stakes Testing, a Struggling School Made a Shocking Choice,” *The New Yorker*, July 21, 2014, pp. 54–65; on Chicago, see Brian A. Jacob and Steven D. Levitt, “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating,” *Quarterly Journal of Economics* 118, no. 3 (August, 2003), pp. 843–77; on “scrubbing” in Cleveland, Ravitch, *Death and Life*, p. 159. Also Goldstein, *Teacher Wars*, p. 227.

13. Goldstein, *Teacher Wars*, pp. 186, 209.

14. Alison Wolf, *Does Education Matter? Similarly*, Donald T. Campbell,

«آزمون‌های موفقیت ممکن است در شرایطی که تدریس معمول با هدف پیشرفت عمومی مدرسه انجام شود شاخص‌های ارزشمندی از موفقیت مدرسه به شمار رود اما هنگامی که نمرات آزمون به هدف برای فرآیند تدریس تبدیل می‌شود، ارزش خود را به عنوان شاخص‌هایی از وضعیت آموزشی از دست می‌دهد و فرآیند آموزش را به روش‌های نامطلوب به انحراف می‌کشاند.»

Mark Palko and Andrew Gelman, “How Schools that Obsess about Standardized Tests Ruin Them as Measures of Success,” *Vox: Policy and Politics*, August 16, 2016, <http://www.vox.com/2016/8/16/12482748/success-academy-schoolsstandardized-tests-metrics-charter>.

۱۵. «ما معتقدیم که این سیستم از تعادل خارج شده است به این معنا که به جای اینکه ابزاری برای پایان دادن به بهترین آموزش ممکن برای همه کودکان باشد، اغلب تلاش برای رسیدن به اهداف تعیین‌شده توسط حکومت تبدیل به هدف شده است. این امر در پدیده‌هایی از جمله تدریس برای آزمون، محدود کردن برنامه درسی و تمرکز بر منابع نامناسب برای دانش‌آموزان حاشیه‌ای است. ما مصرانه از دولت می‌خواهیم که در رویکرد خود برای ایجاد انگیزه در مدارس و تدریس برنامه درسی و شناخت دستاوردهای دانش‌آموزان در یک طیف کامل از برنامه تحصیلی تجدید نظر کند. باید اولویت سیستمی باشد که درباره پیشرفت کودکان به آموزگاران، دانش‌آموزان و والدین اطلاعات دقیق دهد.» (پاراگراف ۸۲)

Select Committee on Children, Schools and Families, *Third Report* (2008). <http://www.publications.parliament.uk/pa/cm200708/cmselect/cmchilsch/169/>

16. Rhodes, *An Education in Politics*, p. 176.

17. Goldstein, *Teacher Wars*, pp. 213–17.

18. Goldstein, *Teacher Wars*, pp. 207–8.

19. Roland Fryer, “Teacher Incentives and Student Achievement: Evidence from New York City Public Schools,” NBER Working Paper No. 16850, March 2011.

20. Goldstein, *Teacher Wars*, pp. 224–26.

21. See the studies cited in Kohn, *Punished by Rewards*, p. 334, fn. 37.

22. Fryer, “Teacher Incentives and Student Achievement,” p. 3.

23. Frederick M. Hess, "Our Achievement-Gap Mania," *National Affairs* (Fall 2011), pp. 113–29.
24. Lauren Musu-Gillette et al., *Status and Trends in the Educational Achievement of Racial and Ethnic Groups 2016* (Washington, D.C.: National Center for Educational Statistics, 2016), p. iv.
25. For a recent confirmation of the Coleman Report's original findings, and of their ongoing relevance, see Stephen L. Morgan and Sol Bee Jung, "Still No Effect of Resources, Even in the New Gilded Age," *Russell Sage Foundation Journal of the Social Sciences* 2, no. 5 (2016), pp. 83–116.
26. Sean F. Reardon, *The Widening Achievement Gap between the Rich and the Poor: New Evidence and Some Possible Explanations* (Russell Sage Foundation, 2012), downloaded from <https://cepa.stanford.edu/content/widening-academic-achievement-gap-between-rich-and-poor-new-evidence-and-possible>.
27. Edward C. Banfield, *The Unheavenly City Revisited* (New York, 1974), pp. 273–74.
28. Among economists, the significance of these qualities has been emphasized by James Heckman, "Schools, Skills, and Synapses," *Economic Inquiry* 46, no. 3 (July 2008), pp. 289–324. Of course, their importance has long been taken for granted by those not wedded to metric fixation. Character qualities of self-control and the ability to defer gratification, however, are themselves linked to cognitive ability, see Richard E. Nisbett et al., "Intelligence: New Findings and Theoretical Developments," *American Psychologist* 67, no. 2 (2012), pp. 130–59, esp. p. 151.
29. Alexandria Neason, "Welcome to Kindergarten. Take This Test. And This One." *Slate*, March 4, 2015.
30. Nisbett et al., "Intelligence," p. 138.
31. Angela Duckworth, "Don't Grade Schools on Grit," *New York Times*, March 27, 2016.
32. Hess, "Our Achievement-Gap Mania," and Wolf, *Does Education Matter?* The declining support for programs for gifted children in Europe is noted in Tom Clynes "How to Raise a Genius: Lessons from a 45-Year Study of Supersmart Children," *Nature* 537, no. 7619 (September 7, 2016).
33. Ravitch, *Death and Life*, passim., and Kenneth Berstein, "Warning from the Trenches: A High School Teacher Tells College Educators What They Can Expect in the Wake of 'No Child Left Behind' and 'Race to the Top,'" *Academe* (January–February 2013), <http://www.aaup.org/article/warningstrenches#.VN62JMZO2AE>; and the powerful testimony of "Teacher of the Year" Anthony J. Mullen, "Teachers Should be Seen and Not Heard," *Education Week*, January 7, 2010, [http://blogs.edweek.org/teachers/teacher\\_of\\_the\\_year/2010/01/teachers\\_shou](http://blogs.edweek.org/teachers/teacher_of_the_year/2010/01/teachers_shou)

## فصل ۹

1. Seana P. Keehan et al., "National Health Expenditure Projections, 2015–2025: Economy, Prices, and Aging Expected to Shape Spending and Enrollment," *Health Affairs* 35, no. 8 (August 2016), pp. 1–10; Atul Gawande, "The Check list," *New Yorker*, December 10, 2007.
2. World Health Report 2000, *Health Systems: Improving Performance*, quoted in Scott

- Atlas, In Excellent Health: Setting the Record Straighton America's Health Care (Stanford, Calif., 2011).
3. Atlas, In Excellent Health.
  4. Ibid., pp. 28–30, 99–105.
  5. Ibid., p. 84; and David M. Cutler, Adriana Lleras-Muney, and Tom Vogl, "Socioeconomic Status and Health: Dimensions and Mechanisms," in Sherry Glied and Peter C. Smith (eds.), *The Oxford Handbook of Health Economics* (New York, 2011), pp. 124–63, esp. 147–53.
  6. Atlas, In Excellent Health, p. 156.
  7. Michael E. Porter and Thomas H. Lee, "The Strategy That Will Fix Health Care," *Harvard Business Review* (October 2013), pp. 50–70, esp. 56.
  8. The discussion of Geisinger is based on Douglas McCarthy, Kimberly Mueller, and Jennifer Wrenn, *Geisinger Health System: Achieving the Potential of Integration through Innovation, Leadership, Measurement, and Incentives* (Commonwealth Fund Case Study, June 2009), and Glenn D. Steele, Jr., "A Proven New Model for Reimbursing Physicians," *Harvard Business Review* (September 15, 2015), by the former CEO of Geisinger.
  9. Peter J. Pronovost et al., "Sustaining Reductions in Central Line-Associated Bloodstream Infections in Michigan Intensive Care Units: A 10-Year Analysis," *American Journal of Medical Quality* 31, no. 3 (2016), pp. 197–202.
  10. Chinitz and Rodwin, "What Passes and Fails as Health Policy and Management," p. 1117.
  11. Patrick Conway, Farzad Mostashari, and Carolyn Clancy, "The Future of Quality Measurement for Improvement and Accountability," *JAMA [Journal of the American Medical Association]* 309, no. 21 (June 5, 2013), pp. 2215–16; James F. Burgess and Andrew Street, "Measuring Organizational Performance," in Sherry Glied and Peter C. Smith (eds.), *The Oxford Handbook of Health Economics* (New York, 2011), pp. 688–706, esp. p. 701; David M. Shahian et al., "Rating the Raters: The Inconsistent Quality of Health Care Performance Measurement," *Annals of Surgery* 264, no. 1 (July 2016), pp. 36–38; J. Matthew Austin, Elizabeth A. McGlynn, and Peter J. Pronovost, "Fostering Transparency in Outcomes, Quality, Safety, and Costs," *JAMA* 316, no. 16 (October 25, 2016), pp. 1661–62.
  12. Chinitz and Rodwin, "What Passes and Fails as Health Policy and Management," p. 1120.
  13. Jason H. Wasfy et al., "Public Reporting in Cardiovascular Medicine: Accountability, Unintended Consequences, and Promise for Improvement," *Circulation* 131, no. 17 (April 28, 2015), pp. 1518–27.
  14. Chinitz and Rodwin, "What Passes and Fails as Health Policy and Management," p. 1118.
  15. N. A. Ketallar et al., "Public Release of Performance Data in Changing the Behaviour of Healthcare Consumers, Professionals or Organisations" *Cochrane Database System Review*, Nov. 9, 2011.

16. Gary Y. Young, Howard Beckman, and Errol Baker, "Financial Incentives, Professional Values and Performance," *Journal of Organizational Behavior* 33 (2012), pp. 964–983.
17. Elaine M. Burns, Chris Pettengell, Thanos Athanasious, and Ara Darzi, "Understanding the Strengths and Weaknesses of Public Reporting of Surgeon-Specific Outcomes," *Health Affairs* 35, no. 3 (March 2016), pp. 415–21, esp. p. 416.
18. D. Blumenthal, E. Malphrus, and J. M. McGinnis (eds.), *Vital Signs: Core Metrics for Health and Health Care Progress* (Washington, 2015), p. 90. On the use of "star ratings" by the National Health Service in England, see Bevan and Hood "What's Measured Is What Matters."
19. Chinitz and Rodwin, "What Passes and Fails as Health Policy and Management," pp. 1114–19.
20. Karen E. Joynt et al., "Public Reporting of Mortality Rates for Hospitalized Medicare Patients and Trends in Mortality for Reported Conditions," *Annals of Internal Medicine*, published online May 31, 2016.
21. M. W. Friedberg et al., "A Methodological Critique of the ProPublica Surgeon Scorecard" (Rand Corporation, Santa Monica, Calif., 2015), <http://www.rand.org/pubs/perspectives/PE170.html>, and David M. Shahian et al., "Rating the Raters: The Inconsistent Quality of Health Care Performance Measurement," *Annals of Surgery* 264, no. 1 (July 2016), pp. 36–38.
22. Cheryl L. Damberg et al., *Measuring Success in Health Care Value-Based Purchasing Programs: Summary and Recommendations* (Rand Corporation, 2014), p. 18. Rachel M. Werner et al., "The Effect of Pay-for-Performance in Hospitals: Lessons for Quality Improvement," *Health Affairs* 30, no. 4 (April 2011), pp. 690–98. Similarly, and most recently, Aaron Mendelson et al., "The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care: A Systematic Review," *Annals of Internal Medicine* 165, no. 5 (March 7, 2017), pp. 341–53.
23. Patricia Ingraham, "Of Pigs in Pokes and Policy Diffusion: Another Look at Pay for Performance," *Public Administration Review* 53 (1993), pp. 348–56; Christopher Hood and Guy Peters, "The Middle Aging of New Public Management: Into the Age of Paradox?" *Journal of Public Administration Research and Theory* 14, no. 3 (2004), pp. 267–82; Chinitz and Rodwin, "What Passes and Fails as Health Policy and Management," pp. 1113–26.
24. Martin Roland and Stephen Campbell, "Successes and Failures of Pay for Performance in the United Kingdom," *New England Journal of Medicine* 370 (May 15, 2014), pp. 1944–49. The phrase "treating to the test" comes from Chinitz and Rodwin, "What Passes and Fails as Health Policy and Management," p. 1115.
25. Burns et al., "Understanding the Strengths and Weaknesses of Public Reporting of Surgeon-Specific Outcomes," p. 418; and Wasfy et al., "Public Reporting in Cardiovascular Medicine"; and K. E. Joynt et al., "Association of Public Reporting for Percutaneous Coronary Intervention with Utilization and Outcomes among Medicare Beneficiaries with Acute Myocardial Infarction," *JAMA* 308, no. 14 (2012), pp. 1460–68; Joel M. Kupfer, "The Morality of Using Mortality as a Financial Incentive: Unintended Consequences and Implications for Acute Hospital Care," *JAMA* 309, no. 21 (June 3, 2013), pp. 2213–14.
26. Richard Lilford and Peter Pronovost, "Using Hospital Mortality Rates to Judge Hospital Performance: A Bad Idea that Just Won't Go Away," *British Medical Journal* (April 10, 2010).

27. Ibid.
28. D. Blumenthal, E. Malphrus, and J. M. McGinnis (eds.), *Vital Signs: CoreMetrics for Health and Health Care Progress* (Washington, D.C., 2015).
29. Ibid., pp. 90–91.
30. Robert Pear, “Shaping Health Policy for Millions, and Still Treating Someone on the Side,” *New York Times*, March 29, 2016.
31. Donald M. Berwick, “The Toxicity of Pay for Performance,” *Quality Management in Health Care* 4, no. 1 (1995), pp. 27–33.
32. Wasfy et al., “Public Reporting in Cardiovascular Medicine”; Claire Noel-Mill and Keith Lind, “Is Observation Status Substituting for Hospital Readmission?” *Health Affairs Blog*, October 28, 2015; as well as <https://www.medicare.gov/hospitalcompare/Data/30-day-measures.html>.
33. David Himmelstein and Steffie Woolhandler, “Quality Improvement: ‘Become Good at Cheating and You Never Need to Become Good at Anything Else’” *Health Affairs Blog*, August 27, 2015.
34. Sabriya Rice, “Medicare Readmission Penalties Create Quality Metrics Stress,” *Modern Healthcare*, August 8, 2015.
35. Shannon Muchmore, “Bill Targets Socio-economic Factors in Hospital Readmissions,” *Modern Healthcare*, May 19, 2016.
36. Michael L. Barnett, John Hsu, and Michael J. McWilliams, “Patient Characteristics and Differences in Hospital Readmission Rates,” *JAMA Internal Medicine* 175, no. 11 (November 2015), pp. 1803–12; and Shannon Muchmore, “Readmissions May Say More about Patients than Care,” *Modern Healthcare* (September 14, 2015).

## فصل ١٠

1. Barry Latzer, *The Rise and Fall of Violent Crime in America* (San Francisco, 2016).
2. On Compstat, see Ken Peak and Emmanuel P. Barthe “Community Policing and CompStat: Merged, or Mutually Exclusive?” *The Police Chief* 76, no. 12 (December 2009); John Eterno and Eli Silverman, *The Crime Numbers Game: Management by Manipulation* (Boca Raton, 2012); Heather MacDonald, “Compstat and Its Enemies,” *City Journal*, February 17, 2010, offers a critique of earlier claims by Eterno and Silverman.
3. Donald T. Campbell, “Assessing the Impact of Planned Social Change” (1976), *Journal of Multidisciplinary Evaluation* (February 2011), p. 34.
4. David Bernstein and Noah Isackson, “The Truth about Chicago’s Crime Rates: Part 2,” *Chicago Magazine*, May 19, 2014.
5. Mac Donald, “Compstat and Its Enemies.”
6. “Police Fix Crime Statistics to Meet Targets, MPs Told,” BBC News, November 19, 2013, <http://www.bbc.com/news/uk-25002927>.
7. Ed Burns interview on “Fresh Air,” National Public Radio, November 22, 2006. See also David Simon interview with Bill Moyers, at [www.pbs.org/moyers/journal/04172009/](http://www.pbs.org/moyers/journal/04172009/)

transcript1.html.

8. Campbell, "Assessing the Impact," p. 35.

## فصل ۱۱

۱. جاناتان شِروِدِن که دانشجوی کالج نیروی دریایی بود نتیجه می‌گیرد که روش‌های معمول ارزیابی کوبین (عملیات ضدشورش) آنقدر ناقص هستند که بهتر است «همه ارزیابی‌های عملیاتی متوقف شوند.»

Jonathan Schroden, "Why Operations Assessments Fail: It's Not Just the Metrics," *Naval War College Review* 64, no. 4 (Autumn 2011), pp. 89–102, esp. 99

2. Connable, *Embracing the Fog of War*, chap. 6.
3. David Kilcullen, *Counterinsurgency* (New York, 2010), p. 2.
4. Ibid., pp. 56–57.
5. Ibid., pp. 58–59.
6. Ibid., p. 60.
7. Connable, *Embracing the Fog of War*, pp. xv, xx.
8. Jan Osborg et al., *Assessing Locally Focused Stability Operations* (Rand Corporation, 2014), p. 9.
9. Connable, *Embracing the Fog of War*, p. 29.

## فصل ۱۲

1. [http://www.simon.rochester.edu/fac/misra/mkt\\_salesforce.pdf](http://www.simon.rochester.edu/fac/misra/mkt_salesforce.pdf).
2. Barry Gruenberg, "The Happy Worker: An Analysis of Educational and Occupational Differences in Determinants of Job Satisfaction," *American Journal of Sociology* 86 (1980), pp. 247–71, esp. pp. 267–68, quoted in Kohn, *Punishment by Rewards*, p. 131.
3. Erik Brynjolfsson and Andrew McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (New York, 2014).
4. Dan Cable and Freck Vermeulen, "Why CEO Pay Should Be 100% Fixed," *Harvard Business Review* (February 23, 2016).
5. Madison Marriage and Aliya Ram, "Two Top Asset Managers Drop Staff Bonuses," *Financial Times*, August 22, 2016.
6. Jeffrey Pfeffer and Robert I. Sutton, "Evidence-Based Management," *Harvard Business Review* (January 2006), pp. 63–74, esp. p. 68.
7. Boris Ewenstein, Bryan Hancock, and Asmus Komm, "Ahead of the Curve: The Future of Performance Management," *McKinsey Quarterly*, no. 2 (2006), pp. 64–73, esp. p. 72.
8. Ewenstein et al., "Ahead of the Curve," pp. 67–68.
9. Tyler Cowen and Alex Tabarrok, *Modern Principles of Macroeconomics*, 3rd ed. (New York, 2014), p. 413.

10. Mark Maremont, "EpiPen Maker Dispenses Outsize Pay," *Wall Street Journal*, September 13, 2016; and Tara Parker-Pope and Rachel RabkinPeachman, "EpiPen Price Rise Sparks Concern for Allergy Sufferers," *New York Times*, August 22, 2016.
11. Matt Levine, "Wells Fargo Opened a Couple Million Fake Accounts," Bloomberg.com, September 9, 2016; and United States of America Consumer Financial Protection Bureau, Administrative Proceeding 2016-CFPB-0015, Consent Order.
12. For other examples, see Gibbons, "Incentives in Organizations," p. 118.
13. Ferraro, Pfeffer, and Sutton, "Economics Language and Assumptions."
14. Douglas H. Frank and Tomasz Obloj, "Firm-Specific Human Capital, Organizational Incentives, and Agency Costs: Evidence from Retail Banking," *Strategic Management Journal* 35 (2014), pp. 1279–301.
15. These examples are cited in *ibid.*, p. 1282.
16. The account that follows draws upon Amar Bhidé, "An Accident Waiting to Happen," *Critical Review* 21, nos. 2–3 (2009), pp. 211–47; and Bhidé, *A Call for Judgment: Sensible Finance for a Dynamic Economy* (New York, 2010), esp. "Introduction"; and Arnold Kling, "The Financial Crisis: Moral Failure or Cognitive Failure?" *Harvard Journal of Law and Public Policy* 33, no. 2 (2010), pp. 507–18, and Arnold Kling, *Specialization and Trade* (Washington, D.C., 2016).
17. Kling, "The Financial Crisis"; and Kling, *Specialization and Trade*, pp. 182–83.
18. Lawrence G. McDonald with Patrick Robinson, *A Colossal Failure of Common Sense: The Inside Story of the Collapse of Lehman Brothers* (New York, 2009), pp. 106–9.
19. Amar Bhidé, "Insiders and Outsiders," *Forbes*, September 24, 2008.
20. The paragraphs that follow draw upon Jerry Z. Muller, "Capitalism and Inequality: What the Right and the Left Get Wrong," *Foreign Affairs* (March–April 2013), pp. 30–51.
21. Hyman P. Minsky, "Uncertainty and the Institutional Structure of Capitalist Economies," *Journal of Economic Issues* 30, no. 2 (June 1996), pp. 357–68; Levy Economics Institute, *Beyond the Minsky Moment* (e-book, April 2012); Alfred Rappaport, *Saving Capitalism from Short-Termism* (New York, 2011).
22. John Asker, Joan Farre-Mensa, and Alexander Ljungqvist, "Corporate Investment and Stock Market Listing: A Puzzle?" *Review of Financial Studies* 28, no. 2 (2015), pp. 342–90.
23. <http://www.businessinsider.com/blackrock-ceo-larry-fink-letter-to-sp-500-ceos-2016-2>.
24. Klarman, *A Margin of Safety*.
25. Nelson P. Repenning and Rebecca M. Henderson, "Making the Numbers? 'Short Termism' and the Puzzle of Only Occasional Disaster," Harvard Business School Working Paper 11–33, 2010. On the negative effects of some pay-for-performance schemes on trust, employee commitment, and institutional productivity, see Michael Beer and Mark D. Cannon, "Promise and Peril in Implementing Pay-for-Performance," *Human Resources Management* 43, no. 1 (Spring 2004), pp. 3–48.
26. Michael C. Jensen, "Paying People to Lie: The Truth about the Budgeting Process," *European Financial Management* 9, no. 3 (2003), pp. 379–406.

27. Gary P. Pisano and Willy C. Shih, "Restoring American Competitiveness," *Harvard Business Review* (July 2009), pp. 11–12.
28. Yves Morieux of Boston Consulting Group, in his TED talk, "How Too Many Rules at Work Keep You from Getting Things Done," July 2015; see also Morieux and Tollman, *Six Simple Rules*.
29. Frank Knight, *Risk, Uncertainty, and Profit* (New York, 1921).
30. Isabell Welpé, "Performance Paradoxon: Erfolg braucht Uneindeutigkeit: Warum es klug ist, sich nicht auf eine Erfolgskennzahl festzulegen," *Wirtschaftswoche* July 31, 2015, p. 88.

## فصل ۱۳

1. Ann Goggins Gregory and Don Howard, "The Nonprofit Starvation Cycle," *Stanford Innovation Review* (Fall 2009); and "The Overhead Myth," [http://overheadmyth.com.b.presscdn.com/wpcontent/uploads/2013/06/GS\\_OverheadMyth\\_Ltr\\_ONLINE.pdf](http://overheadmyth.com.b.presscdn.com/wpcontent/uploads/2013/06/GS_OverheadMyth_Ltr_ONLINE.pdf).
2. See, for example, P. T. Bauer, *Dissent on Development* (Cambridge, Mass., 1976).
3. Mark Moyer, *Aid for Elites: Building Partner Nations and Ending Poverty through Human Capital* (Cambridge, 2016), p. 188. The entire chapter on "Measurement" is invaluable.
4. Andrew Natsios, "The Clash of the Counter-Bureaucracy and Development" (2010), <http://www.cgdev.org/publication/clash-counter-bureaucracy-and-development>; and Natsios, "The Foreign Aid Reform Agenda," *Foreign Service Journal* 86, no. 12 (December 2008), quoted in Moyer, *Aid for Elites*, pp. 188–89.
5. Unnamed USAID official, interviewed by Mark Moyer in 2012, and quoted in Moyer, *Aid for Elites*, p. 190.
6. Moyer, *Aid for Elites*, p. 186.

## فصل ۱۴

1. Moshe Halbertal, *Concealment and Revelation: Esotericism in Jewish Thought and Its Philosophical Implications*, trans. Jackie Feldman (Princeton, 2007), pp. 142–43.
2. Tom Daschle, foreword to Jason Grumet, *City of Rivals: Restoring the Glorious Mess of American Democracy* (New York, 2014), p. x.
3. See on this Jonathan Rauch, "How American Politics Went Insane," *The Atlantic*, July–August, 2016; Jonathan Rauch, "Why Hillary Clinton Needs to be Two-Faced," *New York Times*, October 22, 2016; and Matthew Yglesias, "Against Transparency," *Vox*, September 6, 2016.
4. Cass R. Sunstein, "Output Transparency vs. Input Transparency," August 18, 2016, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2826009](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2826009).
5. Wikipedia, "Chelsea Manning."
6. Christian Stöcker, "Leak at WikiLeaks: A Dispatch Disaster in Six Acts," *Spiegel Online*,

September 1, 2011.

7. Halbertal, *Concealment and Revelation*, p. 164.
8. Joel Brenner, *Glass Houses: Privacy, Secrecy, and Cyber Insecurity in a Transparent World* (New York, 2013), p. 210.

## فصل ١٥

1. Ravitch, *The Death and Life of the Great American School System*, p. 161; Stewart, *The Management Myth*, p. 54.
2. Holmström and Milgrom, "Multitask Principal-Agent Analyses."
3. Merton, "Unanticipated Consequences and Kindred Sociological Ideas: A Personal Gloss," p. 296.
4. Morieux and Tollman, *Six Simple Rules*, pp. 6–16.
5. Lilford and Pronovost, "Using Hospital Mortality Rates to Judge Hospital Performance."
6. Berwick, "The Toxicity of Pay for Performance."
7. George A. Akerlof and Rachel E. Kranton, *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being* (Princeton, 2010), chap. 5, "Identity and the Economics of Organizations."
8. Berwick, "The Toxicity of Pay for Performance."
9. Edmund Phelps, *Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge and Change* (Princeton, 2013), p. 269.
10. Similarly, Scott, *Seeing Like a State*, p. 313.
11. According to Dale Jorgenson of Harvard, the only source of growth of total factor productivity was in IT-producing industries. Dale W. Jorgenson, MunHo, and Jon D. Samuels, "The Outlook for U.S. Economic Growth," in Brink Lindsey (ed.), *Understanding the Growth Slowdown* (Washington, D.C., 2015). On how behavioral metrics in human resources sap initiative, see Lutz, *Car Guys vs. Bean Counters*, pp. ix–x.

## فصل ١٦

1. Young et al., "Financial Incentives, Professional Values and Performance," *Journal of Organizational Behavior* 33 (2012), pp. 964–83, esp. p. 969.
2. Thomas Kochan, commentary on "Promise and Peril in Implementing Pay-for-Performance," *Human Resources Management* 43, no. 1 (Spring 2004), pp. 35–37.
3. J. Matthew Austin, Elizabeth A. McGlynn, and Peter J. Pronovost, "Fostering Transparency in Outcomes, Quality, Safety, and Costs," *JAMA* 316, no. 16 (October 25, 2016), pp. 1661–62.
4. B. S. Frey and M. Osterloh, *Successful Management by Motivation. Balancing Intrinsic and Extrinsic Incentives* (Heidelberg, 2002).
5. Kling, *Specialization and Trade*, p. 33.