

# ارزیابی دانشمندان و آثار علمی آنان

دکتر فاضل لاریجانی  
عضو هیئت علمی وزارت فرهنگ و آموزش عالی  
مجموعه نوروزیان  
عضو هیئت علمی دانشگاه آزاد اسلامی

این مقاله به طور عمده بر آرای «آندره زیابیکی» (Andrzej Ziabicki) مبتنی است که در زمینه ارزیابی کار دانشمندان و فعالیت‌ها و دست‌آوردهای علمی آنان، شیوه‌ای ابتکاری عرضه کرده است، هر چند نواقصی نیز دارد که در بخش انتهایی، با عنوان «نقدی بر مقاله» به آن اشاره خواهد شد. برای توضیح ریاضی برخی مفاهیم این مقاله، بخش کوتاهی با عنوان «پیوست» اضافه شده است که خوانندگان علاقه‌مند در صورت نیاز می‌توانند به آن مراجعه کنند.

### چرا باید دانشمندان را ارزیابی کرد؟

ارزیابی دانش و دانشمندان، برای جوامع صنعتی که در آن‌ها، علم، نه تنها پدیده‌ای فرهنگی و عملی عقلاتی است، بلکه عامل مهمی در رشد اجتماعی و اقتصادی جامعه محسوب می‌شود، امری بدیهی است. ارزیابی کارهای علمی دانشمندانی چون «فیناگورت»، «ارشمیدس»، «لئونارد دواوینچی» و «کارل فون لاین» یا هم عصران آن‌ها ضرورت ندارد، اما در جامعه مدرن، بدون ارزیابی عملی و حمایت مالی، ادامه تحقیقات غیر ممکن است. ارزیابی کارهای دانشمندان و تعیین اعتبار و توان تحقیقاتی آن‌ها، به جایگاه علمی و آموزش در هر کشور بستگی دارد. در لهستان و سایر کشورهای اروپای شرقی، تحقیقات و آموزش، به شدت متمرکز و نهادینه شده‌اند. همه تصمیم‌ها در زمینه حمایت‌های مالی (از محل بودجه دولتی) و استخدام‌ها و ترفیحاتی داخلی، به ارزیابی آن‌ها برمی‌گردد. زندگی یک دانشمند بلند مرتبه، از مقالات، پایان‌نامه‌ها، پرسشنامه‌های تحقیقاتی، حضور در کمیته‌ها، انجمن‌ها و کمیته‌های ارزیابی انباشته شده است. این که همه این بررسی‌ها واقعاً ضرورت دارند یا خیر، موضوع جداگانه‌ای است که در این مقاله به آن نمی‌پردازیم. برخی از ارزیابی‌ها، یقیناً با توسعه روش اعطای درجات و ترفیحات، ساده‌تر شده، کاهش می‌یابند. در حالی که برخی از تصمیم‌گیری‌های مهم، هنوز به ارزیابی نیاز دارند.

کیفیت فرایند ارزیابی از مؤسسه و نظامی به مؤسسه و نظامی دیگر و همچنین از دآوری به داور دیگر فرق می‌کند. لیکن تجدید نظر و همراه کردن چنین ارزیابی با جنبه‌هایی چون بی‌طرفی و مقرون به حقیقت بودن، یکی از موارد با اهمیت این زمان محسوب می‌شود.

### برآورد کیفی ایده‌آل: دقت در مقابل صحت

نتیجه ایده‌آل هر نوع اندازه‌گیری باید درست (یعنی منعکس‌کننده صحیح مقدار مورد اندازه‌گیری) و عینی (یعنی مستقل از شخصی که اندازه می‌گیرد) باشد. در نظریه اندازه (اندازه‌شناسی) از دو خاصیت اصلی بحث می‌شود: صحت و دقت.

صحت مشخصه‌ای از درستی است. این مشخصه، میزان نزدیکی نتیجه حاصل (برآورد) شده را با مقدار واقعی نشان می‌دهد. شرط اولیه صحت، آن است که موضوع مورد اندازه‌گیری، به خوبی تعریف و صورت کمی بیان شده باشد. بیشتر اندازه‌گیری‌ها در

موضوع ارزش‌یابی پژوهش، پژوهشگران و دانشمندان، پدیده‌ای نوظهور و از جمله مسائل عصر حاضر است و شاید بتوان آن را یکی از ویژگی‌ها و مشخصه‌های جوامع صنعتی و مدرن برشمرد، زیرا در این جوامع، علم نه تنها محصولی فکری و فرهنگی است، بلکه بخش جدایی‌ناپذیر توسعه اقتصادی و اجتماعی و از جمله مهم‌ترین مؤلفه‌های آن محسوب می‌گردد. در جوامع صنعتی، تحقیقات هنگامی به مرحله عمل می‌رسند و از پشتیبانی مالی برخوردار می‌شوند که با ارزش‌یابی همراه باشند.

اگرچه تحقیقات علمی صرفاً برای کسب امتیاز صورت نمی‌گیرند و بیشتر چالشی فکری و خلاق برای حل مشکلات به‌شمار می‌روند، اما همواره نحوه ارزش‌یابی دانشمندان و دست‌آوردهای علمی آنان، معضلی اساسی در جریان تحقیقات است.

فعالیت علمی و پژوهشی، فعالیتی چندگانه و دارای ابعاد و جنبه‌های گوناگون است، از این رو به دشواری می‌توان ادعا کرد که شاخص‌ها و شیوه‌های کمی، قابلیت ارزش‌یابی ابعاد کیفی و نوآورانه دست‌آوردهای علمی و پژوهشی را دارا هستند. از این دیدگاه، این پرسش اساسی مطرح می‌شود که «آیا اصولاً علم قابل اندازه‌گیری است» و اگر هست «چگونه»؟

از سوی دیگر، خود فرایند ارزش‌یابی ماهیتاً فرایندی ذهنی است که بر مبنای نوعی دآوری و تشخیص کیفیت‌ها صورت می‌گیرد و به همین دلیل شاید دست‌یابی به نظام عام و فراگیر ارزش‌یابی دشوار باشد، چنان که در حال حاضر هیچ‌گونه شیوه‌ای ثابتی که مورد توافق اکثریت باشد، برای این کار وجود ندارد و هر سازمانی به فراخور نیازها و امکانات خود، به‌طور مستقل، به این کار اقدام می‌کند و البته این موضوع فقط به کشورها محدود نمی‌شود.

نحوه ارزش‌یابی دست‌آوردها و تعیین اعتبار و توان بالقوه تحقیقاتی دانشمندان و پژوهشگران تا حدود زیادی به چگونگی کارکرد علم و آموزش در هر جامعه بستگی دارد.

کیفیت و روش‌شناسی ارزش‌یابی، از یک زمینه و رشته تخصصی، به زمینه تخصصی دیگر، و همچنین از مؤسسه و نهادی به مؤسسه و نهادی دیگر فرق می‌کند.

در این مقاله سعی شده است دو خصیصه اصلی در روش‌شناسی ارزش‌یابی که عبارتست از: صحت و دقت (Reliability & Accuracy)، و اعتبار و عینیت (Objectivity) و نیز روش‌های متداول کمی در «علم‌سنجی» (Scientometrics) بررسی و نقد شود و سرانجام با استفاده از روش‌های ریاضی رویکرد جدیدی عرضه گردد. در این رویکرد ریاضی، ابعاد چندگانه فعالیت دانشمندان و پژوهشگران در قالب فضای دست‌آوردها و اعتبارها (The Space of Achievements & Reputation) به شکل ماتریسی چندبعدي نشان داده شده و هر یک از بردارها و مؤلفه‌های ماتریس نمایانگر یکی از ابعاد کمی و کیفی فعالیت و دست‌آوردهای علمی و پژوهشی است.

فیزیک، شیمی و زیست‌شناسی از این مقوله‌اند، از سوی دیگر، بیشتر تعریف‌ها و ضوابط و معیارها از انسان و علوم انسانی نشأت می‌گیرند.

دقت اندازه‌ای از عینیت و قابلیت بازسازی نتایج است. برای افزایش دقت نیاز است که اندازه‌گیری یک نمونه مشخص چندین بار تکرار و میزان نزدیکی نتایج با یکدیگر مقایسه شود. با این حال، دقت بالا، تضمین‌کننده صحت نتایج نیست؛ زیرا در صورت بروز خطای نظام‌مند در روش اندازه‌گیری، یا انتخاب نامناسب، کمیت مورد اندازه‌گیری یا قاعده‌مند (کالیبره) نشدن ابزارهای اندازه‌گیری، تکرار عمل اندازه‌گیری ممکن است به نتایج بسیار دور از واقع منتهی شود (نظیر خط‌کشی که طول درجات آن متغیر باشد). نتایج ممکن است عینی، دقیق و قابل اندازه‌گیری مجدد باشند، ولی لزوماً صحیح نیستند (نظیر اندازه‌گیری دما با دماسنجی غیراستاندارد).

### ارزیابی دانش و کار دانشمند: کمیت یا کیفیت؟

بسیار خوب می‌بود اگر می‌توانستیم دانش و کار دانشمندان را، با روشی صحیح و دقیق اندازه‌گیری کنیم. همان‌طور که در فیزیک یا شیمی، این کار را انجام می‌دهیم. آیا اندازه‌گیری علم امکان‌پذیر است و آیا صرف وقت برای توسعه روش‌های ارزیابی کمی ارزش دارد؟ در این زمینه دو دیدگاه افراطی وجود دارد.

سنت‌گرایان با معرفی هر نوع فهرست کمیت‌ها مخالفند. آن‌ها معتقدند ارزش کار علمی را نمی‌توان اندازه‌گیری کرد. تهیه فهرست‌های کمیت، مانند تعداد آثار منتشر شده یا میزان مراجعه به کارهای یک دانشمند، به دور از ابهام، انجام‌شدنی نیست. ممکن است تعداد زیادی از کتاب‌ها یا مقالات علمی، دست اول نبوده، خطاهایی داشته باشند. ممکن است مراجعه به این آثار به‌طور تصنعی یا از سوی دوستان و نزدیکان مؤلفان آن‌ها انجام شده باشد. تأثیر دانشمند در پیشبرد علم، ممکن است ناچیز باشد. در حالی که در فهرست کمیت‌ها، چنین مواردی گنجانده می‌شوند. نمونه‌هایی از تبلیغات وسیع و مراجعات مکرر و ارزش‌گذاری به این نوع کارها، بارها در تاریخ علوم دیده شده است. به عکس، تاریخ، دانشمندان بزرگی را به‌خاطر می‌آورد که گرچه تعداد کارهای آن‌ها بسیار محدود است، اما اثری ماندگار در علم برجای گذاشته‌اند: نظیر «آلبرت اینشتین»، «اواریسست گالوا»، ریاضیدان فرانسوی - و «جی. آی. تیلور» - کارشناس ترمودینامیک انگلیسی.

موضوع ارزیابی دانشمندان، ممکن است تحت تأثیر آرای فردی یا حتی دولتی قرار گیرد. در اروپای شرقی، تا سال ۱۹۸۹، بسیاری از ارزیابی‌ها براساس ملاحظات سیاسی و ایدئولوژیک انجام می‌شد. این عوامل ممکن است هنوز هم در بررسی‌ها دخیل باشند. دیدگاه دیگر را کسانی ابراز می‌کنند که اعتقاد دارند با توصیف آماری گروه‌های وسیعی از دانشمندان، می‌توان به روش‌های علمی معمولاً بر مبنای جمع‌بندی ارزش‌ها یا اعتبارهای اختیاری که به کارهای مختلف علمی تخصیص می‌یابند، انجام می‌شوند.

ارزیابی دانش و دانشمندان به دلایل متعدد، باید کیفی و عینی باشد. نخست این که ارزش هر کار دست اول و خلاقانه (در علم، هنر و تکنولوژی) را نمی‌توان با عبارات کمی بیان کرد. دوم این که، چهره دانشمند در آینه کارها و اعتبارات او، موضوعی چندبعدی و شامل عوامل مختلف و سنجش‌ناپذیر است. این عناصر را نمی‌توان با یکدیگر جمع کرد. هر «مقیاس وزنی» از این نوع نیز اختیاری است. سوم این که، مفهوم کیفیت در علم، منحصر به فرد است و به شیوه و هدف ارزیابی بستگی دارد. به همین دلیل فعالیت‌ها و کارهای مختلف یک فرد، ممکن است کم یا زیاد ارزیابی شوند. هنگامی که هدف ارزیابی، استخدام در دانشگاهی یا ارتقا به جایگاهی بالاتر باشد، توانایی‌های تدریس، تجربه دانشگاهی و نگارش کتاب‌های درسی، اهمیت بسیار دارند. همچنین، زمانی که هدف، تخصیص اعتبار به تحقیقات باشد، باید به توانایی تحقیق و نیز کیفیت تحقیقات ناب توجه شود. دشواری‌های تعیین ضوابط ارزیابی، نه تنها در نظام‌های مختلف علمی دیده می‌شوند، بلکه در نظامی واحد نیز همین دشواری‌ها وجود دارند. یک دانشمند، فعالیت‌های مختلفی انجام می‌دهد. کار دانشمندی که داده‌های فراوانی عرضه کرده است و داده‌های او الهام‌بخش نیز نیستند، اما بارها در نوشته‌ها به آن‌ها ارجاع می‌شود در مقابل دانشمند دیگری که یک یا دو کار اساسی منتشر کرده است، ولی بسیار کم به آن‌ها مراجعه شده است، چگونه باید با یکدیگر مقایسه شوند؟ فهرست کمیت‌ها به دانشمند نوع اول بهای بیشتری می‌دهد.

هر چند استدلال‌هایی که ذکر کردیم، با فنون ارزیابی کمی موضوعی مخالفت دارند، اما نمی‌توان استفاده از اطلاعات کمی در فرایند ارزیابی را منکر شد. مجموعه‌ای از فهرست‌های کمی به دقت تعریف شده توأم با کیفیت، بررسی‌های دقیقی را ایجاب می‌کنند و به‌نظر می‌رسد در اختیار داشتن چنین فهرستی، برای یک ارزیابی واقعی ضروری باشد.

این توهم که استفاده از فهرست کمیت‌ها به تنهایی موجب ارزیابی واقعی خواهد شد، خطرناک است؛ زیرا سبب انحراف ذهن از توجه به ضرورت وجود فرایندهای ارزیابی می‌شود.

### فرایند ارزیابی به‌مثابه موضوعی قانونی

ارزیابی دانشمندان و گروه‌های تحقیقاتی، اعطای درجات علمی، جوایز و هدایا و توزیع جوایز تحقیقاتی، بر اساس تصمیم‌گیری و طی روندی قانون‌مند صورت می‌گیرند؛ روندی که در آن، قیاس (مقایسه بین دو موضوع) مداخلت دارد.

شخص یا کمیته ارزیابی، نقش قاضی دادگاه را ایفا می‌کند و تصمیم (یا تشخیص کیفیت) به‌مثابه صدور حکم تلقی می‌شود. تصمیم (یا حکم) به‌مثابه روندی قانون‌مند، مقوله‌ای ذهنی است، اما اختیاری و ارادی نیست. ارزیابی (تعیین حکم) به رأی ارزیاب وابسته است و چنین فرض می‌شود که برای این که فرایند ارزیابی صورت گیرد، لازم است قواعدی (قانون و آیین‌نامه) مهیا باشد. فرایند ارزیابی باید در سه مرحله صورت گیرد. جمع‌آوری

اطلاعات، تجزیه و تحلیل اطلاعات، و تعیین تصمیم‌نهایی.

اطلاعات کامل از یک محقق، طرح تحقیقاتی یا گروه تحقیقاتی، شامل شاخص‌های کمی در فرایند ارزیابی متخصصان آن حوزه است. اهمیت شاخص‌های کمی در فرایند ارزیابی مانند اهمیت مدارک است در دادگاه. متخصصان نیز نقش شاهد را در این دادگاه ایفا می‌کنند و هر دو در فرایند ارزیابی بسیار ضروری هستند.

گرچه در نظر داشتن تمامی اطلاعات مربوط برای ارزیابی ضروری است، اما به معنی اجباری بودن آن تلقی نمی‌شود. وجود اختلاف بین نتایج ارزیابی (تشخیص) قابل قبول است، ولی به توجیه نیاز دارد.

در حالت طبیعی، ممکن است فردی، ارزیابی را همان بررسی دقیق شاخص‌های کمی بداند. دانشمندی که بخواهد کار او مثبت ارزیابی شود (دست‌کم در علوم دقیق)، باید بتواند تعداد مقالات علمی‌ای را که در مجلات معتبر علمی از او منتشر شده، سمینارهای معتبری که در آن‌ها نتایج کارهای علمی خود را عرضه نموده و کارهای علمی‌ای که متخصصان همان حوزه علمی، به آن‌ها رجوع کرده‌اند. برشمارد. بدون شک هیچ معیاری برای تعیین دقیق تعداد هر یک از این موارد وجود ندارد و شاخص‌های کمی را نمی‌توان خودبه‌خود و براساس پیش‌داوری نهایی برآورد کرد، زیرا ممکن است تعداد مقالات دانشمندی محدود باشد، ولی کیفیت علمی آن‌ها به حدی بالا باشد که ارزیابی کار وی به سادگی صورت گیرد.

### یک روش ریاضی: فضای موفقیت و اعتبار

بسیاری از عناصر فعالیت‌های علمی را، با عبارات مناسبی می‌توان توصیف نمود و به صورت شاخص‌های کمی بیان کرد. برای مثال، تعداد آثار منتشر شده در زمینه‌های مختلف، امتیازها، برنامه‌های رایانه‌ای، طرح‌ها و نقشه‌ها، راهنمای پایان‌نامه‌های تحصیلی و ...، اطلاعاتی برای ارزیابی موفقیت‌ها در اختیار ما قرار می‌دهند. افتخارات، جوایز علمی، تقدیرنامه‌ها و عضویت در کمیته‌های علمی، جلوه‌هایی از اعتبارات علمی هستند. در فرایند ارزیابی، شاخص‌های گوناگون و متعددی را باید در نظر گرفت. الگوی کمی پیشرفت یک دانشمند را با ترسیم برداری در فضای  $N$  بعدی می‌توان مشخص کرد. این فضا را، فضای موفقیت و اعتبار می‌نامیم و با  $X$  نمایش می‌دهیم.

$$x \in X \quad (1)$$

عناصر بردار  $x$  متناظر با شاخص‌های کمی، مشخص و آشکار می‌شوند و به خروجی کارهای علمی و اعتبار افراد برمی‌گردند. مثلاً عواملی چون تعداد و تنوع مواردی که فرد منتشر کرده، تعداد پایان‌نامه‌هایی که با راهنمایی او تهیه و تدوین گردیده، یا میزان ارجاعاتی که به کارهای وی شده است، می‌توانند عناصر بردار  $x$  را تشکیل دهند. بعد فضای  $x$  و همچنین فهرست شاخص‌های یاد شده ( $x_i$ ) (ها) به اهداف ارزیابی وابسته است. فضای موفقیت و اعتبار، خطی است،

ولی متریک (متری) نیست.

خطی بودن به این معناست که عناصر (شاخص‌های) یکسان ( $x_i$ ) مربوط به یک دانشمند یا گروهی از دانشمندان یا مؤسسات در دوره‌های مختلف، یا  $x$  مربوط به هر یک از اعضای گروه دانشمندان را می‌توان با هم جمع کرد. همچنین برای کارهایی که چند نویسنده عرضه کرده‌اند (نظیر کتاب‌ها، مقالات و حق ثبت‌ها) باید از عمل جمع در نظریه مجموعه‌ها (اجتماع،  $U$ ) استفاده کرد.

غیرمتریک بودن فضای  $x$  از این حقیقت ناشی می‌شود که عناصر مختلف فضا را نمی‌توان با یکدیگر مقایسه کرد، چون هر عضو دارای یک وارون (عضو متقابل) نیست. متریک نبودن، مانع از برآورد جامع کیفیت و مقایسه دانشمندی با سیمای فعالیتی مختلف می‌شود. دانشمند  $A$  را که دارای تعداد زیادی اثر چاپ شده مرجع ولی کمیاب است، نمی‌توان از نظر کیفی، با دانش  $B$  که دارای تعداد محدودی مقاله قابل دسترس برای همگان است، مقایسه کرد. غیرمتریک بودن فضای  $x$  در مقایسه یک دانشمند محقق (به لحاظ تعداد آثار منتشر شده، بدون آموزه‌های تحصیلی) با یک پروفیسور دانشگاهی که مهم‌ترین موفقیت‌های او به امور دانشگاهی و آموزشی وابسته است (نظیر تدریس دروس، هدایت دانشجویان کارشناسی ارشد و دکترا) به صورت کاملاً بدیهی مشهود است.

### استفاده از شاخص‌های کمی در فرایند ارزیابی

آن دسته از شاخص‌های کمی که موفقیت‌ها و اعتبارات علمی را مشخص می‌کنند، بخش مهمی از اطلاعات (شواهد و مدارک) را در ارزیابی به خود اختصاص می‌دهند. راه‌های بسیاری برای استفاده از این اطلاعات وجود دارد. انتخاب شاخصی مفرد (نظیر تعداد آثار منتشر شده یک فرد)، می‌تواند به قضاوت یک سوئیه بسیاری از عوامل مهم دیگر منجر شود. ارزیابی چندبُعدی، اغلب بر مبنای شاخص‌هایی کمی استوار است که به مشخصه عددی معینی به نام برآورد جامع (GE)  $Global Estimation$  تبدیل شده‌اند. در روش دیگری که پیشنهاد می‌کنیم، اطلاعات مربوط به عناصر گوناگون موفقیت علمی، به‌طور موازی به شکل بُرداری در فضای  $x$  نمایش داده می‌شوند.

### برآوردهای جامع

با وجود این که نمایه فعالیت‌های علمی، چند وجهی است و شاخص‌های مختلف را با یکدیگر نمی‌توان مقایسه کرد، لیکن محاسبه برآورد عددی کیفیت علمی، کاری کاملاً معمولی است.

قدیمی‌ترین برآوردهای جامع براساس شاخصی منحصر به فرد، نظیر تعداد آثار منتشر شده یا تعداد ارجاعات، استوار شده‌اند. این موضوع در فضای  $x$  به این معنی است که همه مؤلفه‌ها (شاخص‌ها)، جز یکی (مثلاً مؤلفه  $x_0$ ) را جدا کنیم و آن را مبنای ارزیابی کمی قرار دهیم.

$$GE = x_0 \quad (2)$$

مبنا قرار دادن تعداد آثار منتشر شده، بدون در نظر گرفتن نوع و

درجه اهمیت آن‌ها، در روش‌های قدیم ارزیابی دانشمندان و مؤسسات علمی مرسوم بود. در چنین روشی، هر ورودی در فهرست آثار منتشر شده، به صورت یک واحد، شمارش می‌شد: مثلاً تکننگاری، انتشار مقاله‌ای تحقیقی در مجله علمی معتبر، درج یادداشتی کوتاه در مجله‌ای داخلی، همگی به یکسان محاسبه می‌شدند. حال آن‌که در نظر داشتن تقدم و تأخر آثار منتشر شده به لحاظ میزان اهمیت، ارزیابی را عمیق‌تر می‌کند. یکی از راه‌حل‌های پیشنهادی، این است که آثار منتشر شده مختلف را به رده‌هایی طبقه‌بندی کنیم و به هر یک وزن یا اعتبار ( $s_i$ ) مناسبی، تخصیص دهیم که بیانگر ارزش نسبی هر رده باشد. اگر  $x_i$  نشان‌دهنده تعداد آثار منتشر شده فرد در رده  $i$ ام باشد، خواهیم داشت:

$$GE = \sum_{i=1}^k s_i x_i \quad (3)$$

ضریب  $s_i$  معمولاً به روش‌های دلخواه در نظر گرفته می‌شود (مثلاً تکننگاری ۲۰ امتیاز، مقاله تحقیقی دست اول ۳ امتیاز، مقاله چاپ شده در نشریه‌ای محلی ۱ امتیاز و...). همین روش را می‌توان برای سایر مؤلفه‌های بردار  $x$ ، مانند تقدیرنامه‌ها، سخنرانی‌ها، فعالیت‌های دانشگاهی و غیره نیز به کار برد. این شیوه امتیازهای  $s_i$  مقیاس محاسبه را تشکیل می‌دهد. همه مقیاس‌های از این نوع، دلخواه هستند. هیچ روش قانع‌کننده‌ای وجود ندارد که به کمک آن بتوان ارزش آثار منتشر شده را با تقدیرنامه‌ها یا سایر فعالیت‌های علمی مقایسه کرد. دلخواه بودن مقیاس‌های محاسبه، برآورد جامع رابطه (۳) را، نه تنها غیر دقیق، بلکه غیر قابل ارائه می‌کند. مقیاس‌های مختلف و عمومیت نداشتن آن‌ها از دیدگاه ارزیاب‌ها، باعث می‌شود که مجموعه معینی از  $x_i$ ها، به روش‌های مختلف ارزیابی شود.

یک راه حل، در نظر گرفتن «عوامل وزنی» است، یعنی فقط مقالاتی را در نظر بگیریم که در مجلات علمی منتشر شده‌اند. این کار در نشریه ضمیمه ارجاع علمی (SCI) به چاپ رسیده است. عامل وزنی برای آلمین مجله را  $f_i$  می‌نامیم که برابر است با فراوانی ارجاع به مقاله‌ای متوسط در این مجله، براساس این عامل‌های وزنی، هر چند برآورد جامعی قابل ارائه است، ولی دقیق نیست.

$$GE = \sum_{i=1}^k f_i x_i \quad (4)$$

گرچه عامل‌های وزنی، اطلاعات جالبی درباره مجلات و سمت و سوی آثار منتشر شده در اختیار می‌گذارند، ولی نمی‌توانند ابزار مناسبی برای اندازه‌گیری ارزش مقالات به کار گرفته شده باشند. فراوانی مراجعات، کاملاً به نظم انتشار، زبان و نوع مجله بستگی دارد (برای مثال، مجلاتی که تحقیقات انجام شده را مرور می‌کنند در مقابل مجلاتی قرار می‌گیرند که تحقیقات دست اول را منتشر می‌نمایند، حجم مجله نیز در میزان مراجعه مؤثر است. همچنین مجلاتی تخصصی در مقابل مجلات عمومی یا مجلاتی هستند که مربوط به چند رشته علمی مختلفند). جدول ۱، عامل‌های وزنی مربوط به مجلات علمی را که به آن‌ها بسیار ارجاع می‌شود (رتبه ۱)، در موضوعات علمی مختلف نشان می‌دهد.

بدیهی است که عامل‌های وزنی مجلات رتبه ۱ در شاخه‌های علمی مختلف، در یک ضریب ۵۰، با هم فرق دارند. اگر ارزیابی بر اساس تعداد آثار منتشر شده صورت گیرد در عامل وزنی ضرب می‌شود. برآورد جامع برای دانشمندی که به تعداد یکسان مقاله دارند، براساس میزان ارجاع مقالات چاپ شده در مجلات معتبر علمی به مقالات این دانشمندان، انجام می‌شود. ارزیابی نسبی دانشمندان (یا مؤسسات) در عرصه‌های مختلف، موضوعی علمی - دانشگاهی تلقی نمی‌شود؛ زیرا به درستی معلوم نیست با آن دسته از کارهای چاپ شده در مجلات علمی داخلی که در SCI به آن‌ها اشاره نشده (و اتفاقاً تعداد آن‌ها نیز زیاد است) چه باید کرد. این موضوع، به ویژه تمام مجله‌هایی را که به زبان‌هایی غیر از انگلیسی، فرانسه، آلمانی و تاحدی روسی منتشر می‌شوند، دربرمی‌گیرد. آیا مقاله‌هایی که در چنین نشریاتی منتشر می‌شوند، باید به کلی از نظر دور بمانند؟ به این ترتیب، عدم کفایت برآورد جامعی که از این نوع ارزیابی، مبتنی بر عامل‌های وزنی صورت می‌گیرد، بدیهی به نظر می‌رسد. بدون شک این پیشنهاد که برای ارزیابی باید سیاست ثابتی اتخاذ شود، ضروری است و شاید سیاست اتکا بر عامل‌های وزنی بتواند در میزان تولید آثار مختلف، تغییرات مطلوبی به وجود آورد، اما باید توجه داشت که هدف پژوهش، حل مسائل است نه کسب افتخارات (یا عامل‌های وزنی). جایگاه آثار منتشر شده، باید به جای توجه به عامل وزنی، از طریق محتوا و اندیشه مرتبط با آن، محاسبه شود. اگر ارزیابی براساس عامل‌های وزنی به مثابه قاعده‌ای پذیرفته شده، معمول گردد، آن‌گاه لازم است دانشمندان به جای پرداختن به تحقیقات ناب، به نوشتن مطالب تکراری روی آورند و در این صورت اهمیت مطالب آن‌ها تنها در این است که مروری است بر کارهای پیشین. همچنین ارزیابی براساس تعداد آثار منتشر شده، اغلب باعث می‌شود که یک کار علمی به تعداد کار کوچک‌تر تقسیم شود. در اروپای شرقی، این عامل، یعنی توجه به تعداد آثار انتشار یافته، باعث شده است دانشمندان آثار خود را به جای مجلات محلی و به زبان ملی، در مجلات معتبر بین‌المللی به چاپ برسانند، بنابراین به نظر می‌رسد عامل‌های وزنی بهترین ملاک برای ارزیابی نباشند.

با مقایسه مجله‌های مختلف در یک زمینه علمی (جدول ۲)، ممکن است این تصور ایجاد شود که عامل‌های وزنی، به حجم نسبی و تخصصی بودن آثار منتشر شده وابسته‌اند.

بیشتر مجله‌هایی که عامل وزنی بالا دارند. از نوع مجلات مروری‌اند و انتخاب‌های مؤلف محسوب می‌شوند. شیمی دان می‌توان مقاله مروری خود را به مجله **Chemical Reviews** (رتبه ۱،  $f = 10/40$ ) که عامل وزنی بالایی دارد، عرضه کند. اما وقتی تحقیق دست اول او، به شکل گزارش منتشر شود، باید به انتشار مقاله‌اش در مجله‌ای بسیار سطح پایین تر قناعت کند. اگر حوزه کاری او، شیمی کاربردی باشد، مجله **پرمراجع Angewandte Chemie International Edition** (رتبه ۳،  $f = 5/103$ ) در دسترس خواهد بود. برای مطالعه درباره واکنش‌های شیمیایی حالت جامد، انتخاب طبیعی، مجله **Journal of Physics and Chemistry Solids**

## جدول ۱: عوامل وزنی تعدادی از مجلات معتبر در چند شاخه علمی

شاخه علمی	رتبه	مجله	عامل وزنی (f)
بیوشیمی و بیولوژی مولکولی	۱	Annual Review of Biochemistry	۴۸/۳۱۳
ایمونولوژی	۱	Annual Review of Immunology	۲۵/۴۰۸
پزشکی	۱	New England Journal of Medicine	۲۱/۱۴۸
فیزیک	۱	Review of Modern Physics	۱۵/۱۲۸
شیمی	۱	Chemical Review	۱۰/۴۰۰
مکانیک	۱	Annual Review of Fluid Mechanics	۴/۱۷۱
ریاضیات	۱	Publications Mathmatiques (Paris)	۳/۱۵۴
علم پلیمر	۱	Macromolecules	۲/۳۹۵
دیرین‌شناسی	۱	Paleobiology	۴۱/۵۷
کشاورزی	۱	Journal of Agricultural & Food Chemistry	۱/۱۶۵
انرژی	۱	Energy & Fuels	۱/۴۶۱
جنگل‌داری	۱	Holzforschung	۰/۹۶۶
مهندسی	۱	Journal of Research of the National Bureau of Standards	۰/۹۳۸

خواهد بود که رتبه آن ۲۶ و عامل وزنی آن، فقط ۰/۹۱۶ است. در سایر شاخه‌ها نیز، موقعیت‌های مشابهی وجود دارد. بدیهی است بسیاری از مجلات معتبر بین‌المللی که اعتبار زیاد دارند، دارای عامل وزنی نسبتاً کمی هستند. در فیزیک، بیشترین ارجاع به مجله **Review of Modern Physics** صورت می‌گیرد که عامل وزنی (f) آن برابر است با ۱۵/۱۲۸، اما بیشتر مجله‌های معتبر همچون **Proceedings of Royal Society, Journal de Physique, Physica, Journal of Physics** عامل‌های وزنی کمتر از ۱۰ دارند. آیا انتشار مقاله در این مجلات، به دلیل «ضرورت» اجتناب‌ناپذیر است؟

این وضعیت تا حدودی در بیوشیمی و بیوفیزیک متفاوت است. در دهه گذشته، علوم زیستی عمومیت بسیار یافته و بهترین حمایت‌ها را جلب کرده‌اند. این موضوع را می‌توان به وضوح در تعداد دانشمندان محقق در این زمینه‌ها و فراوانی ارجاعاتی که به مقالات منتشر شده در مجلات بیوشیمی می‌شود، ملاحظه نمود. در این جا، عامل‌های وزنی بالا، نه تنها به مجلات مروری (شماره‌های ۱ و ۲ در فهرست) تعلق دارند، بلکه طیف وسیعی از مجلات تخصصی را نیز که مقالات اصلی گزارشی را چاپ می‌کنند، دربرمی‌گیرند. آیا این موضوع، به این معنی است که همه زیست‌شناسان، به علت عامل وزنی بالا، باید به سوی زیست‌شناسی مولکولی و بیوشیمی گرایش پیدا کنند؟ کاربرد مکانیکی عامل‌های وزنی در ارزیابی دانشمندان، ممکن است به تعصب و انعطاف‌ناپذیری جدی و بروز وجوه نامطلوبی از تحقیقات و انتشار نتایج آن‌ها منجر شود.

### مشخصه‌ای چندبُعدی: بُردار موفقیت و اعتبار

بدیهی است که صفات و مشخصات علمی، چندبُعدی هستند و

نمی‌توانند با مشخصه واحد و فراگیری بیان شوند. مشخصه‌های فردی، قابل مقایسه نیستند. تعداد کل آثار منتشر شده، یا تعداد آثار منتشر شده‌ای که به وسیله مضارب یا عامل‌های وزنی، وزن‌دار شده‌اند، شاخص‌هایی برای تشخیص یک دانشمند نیستند. تحقیق دست‌اولی، که در مجله معتبری به چاپ می‌رسد، با یک دوجین یا صد مقاله دست دوم که در مجله‌ای محلی یا گزارش داخلی منتشر شده‌اند، هم‌تراز نیست. موفقیت‌های دانشگاهی (کتاب‌های درسی، دروه‌ها، تدریس در دوره دکترا) را، نمی‌توان با نتایج تحقیقات یا فعالیت‌های گروهی مقایسه کرد. بنابراین، اطلاعات کامل درباره دانشمند یا گروه تحقیقاتی خاص، باید به‌طور موازی، شامل همه صفات قابل تصور باشد، نه این‌که مجموع یا میانگین وزنی در نظر گرفته شود. این موضوع، اصلی‌ترین اختلاف میان روش ما و روش‌های کمی ارزیابی است که در آن‌ها از اعتبار، مقیاس‌های مسئولیت یا فهرست‌های جامع استفاده می‌شوند.

یک راه برای برآورد جامع (یک‌بُعدی یا چندبُعدی) متکی بر مقیاس‌های دلخواه مسئولیت، استفاده از اطلاعات موازی چندبُعدی درباره فعالیت‌های مختلف علمی است. به بیان ریاضی، چنین مشخصه‌ای با بُرداری در فضای موفقیت و اعتبار مشخص می‌شود. N مؤلفه چنین بُرداری را، اعدادی تشکیل می‌دهند که بیانگر مهم‌ترین شکل‌های تحقیقات و خروجی‌های دانشگاهی و همچنین اعتبار علمی هستند. مؤلفه‌های بُردار x شامل شمارش روی داده‌های مختلف است و هیچ مضرب دلخواه یا نامربوط (مانند عوامل وزنی) را شامل نمی‌شود. بُردار موفقیت و اعتبار، شامل اطلاعات معینی است و مبنای کاملی برای ارزیابی در اختیار نمی‌گذارد.

تعداد و فهرست عناصر مربوط، به هدف ارزیابی بستگی دارد.

جدول ۲: عوامل وزنی انتخاب در درون شاخه‌های علمی

عامل وزنی (f)	مجله	رتبه
		شیمی
۱۰/۲۰۰	Chemical Review	۱
۶/۶۶۱	Accounts of Chemical Research	۲
۲۵/۱۰۳	Angewandte Chemie, International Edition	۳
۴/۵۶۶	Journal of the American Chemical Society	۴
۱/۸۸۶	Recueil des Travaux Chimiques de Pays - Bas	۱۲
۰/۹۱۶	Journal of Physical & Chemical Solids	۲۶
		فیزیک
۱۵/۱۲۸	Review of Modern Physics	۱
۸/۲۳۴	Physical Reports	۲
۸/۲۱۳	Physical Review Letters	۳
۲/۲۶۷	Annals of Physics (New York)	۱۱
۱/۹۲۸	Journal of Physics, A	۱۴
۱/۴۱۳	Proceedings of the Royal Society (London), A	۲۰
۱/۲۳۵	Physical A	۲۳
۰/۹۲۴	Journal de Physique (Paris)	۲۹
		مکانیک
۴/۱۷۱	Annual Review of Fluid Mechanics	۱
۱/۶۰۲	Journal of Fluid Mechanics	۲
۱/۵۴۴	Journal of Non - Newtonian Fluid Mechanics	۳
۱/۰۷۰	Journal of Rheology	۵
۱/۰۳۱	Archive for Rational Mechanics Analysis	۵
۰/۶۱۷	Rheologica Acta	۱۳
۰/۱۸۸	Zeitschrift für Angewandte Mathematik und Mechanik	۳۵
		بیوشیمی و بیولوژی مولکولی
۴۸/۳۱۳	Annual Review of Biochemistry	۱
۲۴/۲۱۲	Annual Review of Cell Biology	۲
۲۳/۹۱۳	Cell	۳
۱۸/۱۶۷	Advances in Cyclic Nucleotide Protein Research	۴
۱۶/۵۰۰	Advances in Protein Chemistry	۵
۲/۲۹۱	Archives of Biochemistry and Biophysics	۴۷
۲/۲۷۸	Biochim et Biophysica Acta	۴۸
۲/۱۱۵	Biopolymers	۵۴
۲/۰۰۰	Journal of Bioenergetics and Biomembranes	۵۹
۱/۰۱۶	International Journal of Biological Macromolecules	۹۸

این فهرست، می‌تواند در گروه‌های مختلف علمی و هنری متفاوت باشد. در مقایسه با جریان حقوقی، فهرست مدارک مورد نیاز (شواهد) از طریق مقررات فرایند قضاوت، تعیین می‌شود. صفات کمی به کار رفته، در محاسبه برآورد جامع معادله (۳)، باید در معادله (۵) نیز ظاهر شوند.

(۵)

$x_1$	تعداد منتشر شده رده اول
$x_2$	تعداد آثار منتشر شده رده دوم
-----	-----
$x_j$	تعداد آثار منتشر شده رده $j$ ام
-----	-----
$x_x$	تعداد ارجاعات
$x_x + 1$	تعداد مقالات عرضه شده در همایش رده اول
$x_x + 2$	تعداد مقالات عرضه شده در همایش رده دوم
-----	-----
$x_m$	تعداد مقالات عرضه شده در همایش رده $m$ ام
$x_m + 1$	تعداد نزهایی که هدایت کرده است
$x_m + 2$	تعداد کتابهای درسی
$x_m + 3$	تعداد جایگاه‌های کمیته‌ای
$x_m + 4$	تعداد عضویت‌ها در بردهای سردبیری
-----	-----
-----	-----
$x_n$	-----

در مؤسسه چندمنظوره تحقیقات بنیادین تکنولوژی در دانشگاه علوم لهستان، برای ارزیابی دانشمندان محقق و واحدهای تحقیقاتی، از مجموعه‌ای حاوی ۳۰ مؤلفه استفاده می‌شود. این مؤلفه‌ها، شامل هفت رده از آثار منتشر شده، حق ثبت‌ها، بنیادها، طراحی‌ها، ارجاعات، صفات مشخصه دانشگاهی و اعتبارات و فعالیت‌های گروهی است.

صفاتی که در بُردار  $x$  به کار می‌روند، باید خوب تعریف شده باشند (خوش تعریف باشند). ما به جای عامل‌های وزنی، آثار منتشر شده مجزا در رده‌های مختلف را پیشنهاد می‌کنیم. برای اجتناب از هر اشتباهی، باید هر رده به روشنی تعریف شود. لازم است تکنگاشت‌ها، که تلفیق جامعی از مبحث ویژه به روز شده (روزآمد) را عرضه می‌کنند، از مقالات مروری که حاوی نتایج مطالعات نویسندگان مختلف هستند، یا از مقالات تحقیقاتی دست اولی که آن‌ها را مؤلف به دست آورده است، جدا شوند. مقاله‌های عمومی و کتاب‌های درسی، رده‌های دیگری را تشکیل می‌دهند. با توجه به این که بسیاری از مجلات علمی فقط مقاله‌هایی از نوع خاص را منتشر می‌کنند، معرفی رده‌های مختلف آسان می‌شود. گاهی مجلات فقط برخی مقالات تحقیقی دست اول، بعضی مقالات عمومی یا تکنگاشت‌های مروری منتشر می‌کنند. گاهی اوقات خود

مجله (یا ناشر)، مشخصه و معیار آثار منتشر شده را نشان می‌دهد و رده آن‌ها را مشخص می‌کند. مقاله‌های منتشر شده در مجله‌های محلی (یا کتاب‌هایی که به صورت محلی چاپ می‌شوند) باید از مقالاتی که در مجلات معتبر بین‌المللی به چاپ می‌رسند یا مؤسسات انتشاراتی معتبر جهانی آن‌ها را منتشر می‌کنند، تفکیک گردند. در هر طبقه، دانشمندان از جایگاه مجله یا ناشر مطلع هستند. به این ترتیب، این موضوع که تعیین رده آثار منتشر شده را داوری بی‌طرف، بررسی و تصحیح کند، دارای اهمیت است. گرچه SCI می‌تواند در این زمینه نقش مفیدی ایفا کند، ولی رتبه‌ها و عوامل وزنی را نمی‌توان خودبه‌خود از این منبع اخذ کرد. برخی سازمان‌های امتیازدهنده (همچون بنیاد بین‌المللی علوم که تحقیقات در اتحاد شوروی سابق را حمایت می‌کرد)، دارای فهرست ویژه‌ای از مجلات خوب بودند که در SCI به آن‌ها اشاره‌ای نشده است.

سایر عناصر تحقیقات، نظیر حق ثبت‌ها، کشفیات دست اول، طراحی‌ها و برنامه‌های رایانه‌ای که جداگانه شمارش می‌شوند نیز اهمیت دارند، مخصوصاً در سازمان‌هایی که تحقیقات محض و کاربردی را به کار می‌گیرند. این روش برای موفقیت‌های دانشگاهی هم استفاده می‌شود.

یکی از مشخصات عمومی اعتبار عبارت است از تعداد ارجاعات در ادبیات علمی که در SCI یا مراکز مشابه بیان می‌شود. فراوانی ارجاعات، فقط به رتبه مجله و عمومیت مبحث و سبک و مشخصه آثار منتشر شده بستگی دارد. معمولاً به کتاب‌ها، تکنگاشت‌ها و مقالات مروری، بیش از تحقیقات دست اول مراجعه می‌شود. به مقالات تجربی (یا حاوی آزمایش) که شامل داده‌های فیزیکی و شیمیایی جدید هستند نیز، نسبت به مقالات نظری، بیشتر ارجاع صورت می‌گیرد. به طعنه می‌توان گفت، بیشترین ارجاع مربوط به مقالاتی است که دارای خطا هستند. خوانندگان از این که به خطاهای دیگران اشاره کنند، شاد می‌شوند. این خطاها، معمولاً شامل موارد مثبتی هستند که در ارزیابی اعتبار، از آن‌ها استفاده شده است.

سایر موارد اعتباری، شامل احکام و افتخارات، دعوت برای عرضه مقالات در سمینارهای علمی، عضویت در جوامع دانشمندان، انجمن‌های تخصصی، هیئت سردبیری مجلات علمی، انجمن‌های علمی و... است. برخی از این موارد (به ویژه دعوت‌ها و عضویت در کمیته‌ها)، ممکن است براساس روابط سیاسی و اجتماعی (نظیر هم‌ترازی جغرافیایی در سازمان‌های بین‌المللی) استوار باشند. برآورد اعتبار براساس عضویت در کمیته‌ها، حضور در مجامع و غیره، به طور ذاتی و ماهیتاً، با انتشار مقالات تحقیقاتی دست اول و افتخارات دانشگاهی متفاوت است.

### بررسی‌های دقیق (موشکافانه)

اطلاعات اساسی دیگری که در فرایند ارزیابی از آن‌ها استفاده می‌شود، عقیده و نظر متخصصان و کارشناسان، در آن زمینه است. در حال حاضر، بررسی‌های دقیق، پایه‌های اصلی بررسی محسوب



می شوند. وظیفه اصلی بررسی های دقیق، توصیف کمی دانشمند در حال بررسی و پاسخ به پرسش هایی است که هیئت ارزیابی کننده مطرح می کند. بررسی های دقیق، باید اطلاعات را که در فهرست کمیته ها وجود دارند، مشخص و تفسیر کنند.

در انتخاب ارزیاب ها، باید به صلاحیت، تمایلات احتمالی و اختلاف نظرهای ممکن و علایق آن ها توجه شود. در پرسشنامه های بنیاد بین المللی علوم، ارزیاب ها موظفند در زمینه اختلاف نظرهای احتمالی، مواردی را در نظر گیرند و لازم است این موضوع، تشویق و حمایت شود. بدون شک این کار موجب حذف همه خطاها نخواهد شد، ولی شاید عقاید واقعی تری را در این امر مشارکت دهد. در موقعیتی حقوقی، تداخل عقاید، معمولاً در فرایند بازپرسی یا پرسش و مقابله (روبه روستازی) نشان داده می شود. در لهستان، این کار را شوراهای دانشکده ها که درجات علمی را اعطا می کنند و همچنین کمیته های مرکزی اعطای درجات و عناوین علمی که همه درجات علمی دکترا و تخصص های نامزدهای دریافت این عناوین را تصویب می کنند، برعهده دارند در سایر فرایندهای ارزیابی، بررسی های دقیق، فقط از طریق فرم های پر شده و تجزیه و تحلیل آن ها به کمک هیئت ارزیابی، در غیاب ارزیاب ها انجام می شود. هیئت ارزیابی، مسئول استفاده انتقادی از بررسی های دقیق است.

برخی داوران، فاقد صلاحیت یا بهانه گیر هستند و می کوشند عقاید واقعی خود را بروز ندهند. ناکارایی بررسی، به عوامل متعدد و همچنین به نحوه تنظیم پرسش هایی که برای داوری تهیه شده اند، بستگی دارد. سؤالات هوشمندانه ای چون: «کار اصلی دانشمند در حال داوری، در چه زمینه ای است؟» «چه برخوردی را با تحقیق عرضه شده مناسب می دانید؟» «چگونه کار دانشمند در حال داوری را با زمینه کاری او (کارهای اخیر او در این زمینه) مقایسه می کنید؟» به جای پرسش هایی چون: «آیا دانشمند مورد نظر، تحقیقات خوبی انجام داده است (مقالات خوبی منتشر کرده است)؟»، «آیا پیشنهاد او حاوی عناصر قابل اندازه گیری جدید است؟»، سرچشمه ای از اطلاعات در اختیار می گذارند. روش معمول تقاضا از داوران، برای برآورد عددی کردن کیفیت، باید کنار گذاشته شود. هر داور، برای خود درجه بندی خاصی از مقادیر، تقدم ها و پیش داوری ها دارد و برآورد یا طبقه بندی او دارای معنی محدودی است. هنگامی که چند دانشمند را داوران مختلف ارزیابی می کنند، میانگین گرفتن از نمره های داوران نیز، کمک زیادی نمی کند. البته بر مقایسه چندین پرسشنامه یا بررسی های توصیفی، تأکید می کنیم. برای تجزیه و تحلیل دقیق، باید به واگرایی شدید عقاید توجه کرد و در صورت لزوم، چند داور جدید نیز باید به جمع داوران افزوده شوند.

### داوران به جای هیئت ارزیابی

درک نقشی که داوران (شاهدان) ایفا می کنند، از یک سو و اعضای کمیته ارزیابی (دادگاه) از سوی دیگر، قوانینی بر ارتباط متقابل و رفتار مشترک در فرایند ارزیابی حاکم می کند.

نقش تصمیم گیرندگان و داوران با هم تفاوت دارد. این نقش ها نباید با یکدیگر تداخل داشته باشند. به ویژه، اعضای کمیته ارزیابی، نباید داوری کنند. آن ها باید عقاید خود را بر اساس اطلاعات به دست آمده استوار کنند و در مقابل داوران و دانشمندان در حال داوری، موضع مستقلی داشته باشند. البته این اصل بدیهی، همواره در ارزیابی ها رعایت نمی شود.

برای بررسی واقعی، داور باید امین و در کار خود متخصص باشد. تشکل داوران، بخشی از جامعه علمی را تشکیل می دهد. آن ها به مثابه متخصص، در برخی موارد ارزیابی شرکت می جویند و داوران و ارزیاب های دیگر درباره آن ها داوری می کنند. به طور کلی برتری دادن به محشی ویژه یا تحقیقی خاص، امری طبیعی است، اما باید در فرایند داوری در نظر داشت که انتخاب شیوه بررسی و ارزیابی، در صلاحیت و حوزه اختیارات کمیته ارزیابی را برای اتخاذ روشی انتقادی در بررسی ای دقیق، افزایش می دهد. وجود تعصب در بررسی ها را، با افزایش تعداد داوران می توان برطرف کرد.

از جمله شرایط اصلی برای اعضای کمیته ارزیابی (تصمیم گیرندگان)، به بی طرفی، وسعت نظر و درک کافی از پهنه های وسیع علمی و تجربه و آزمودگی در بررسی کار دانشمندان در رده های مختلف می توان اشاره کرد. داشتن دانش لازم از موضوعی که مورد ارزیابی است، کاملاً ضروری است. همه اطلاعات مرتبط با موضوع، باید در فهرست کمیته ها و بررسی دقیق مدنظر قرار گیرند. آنچه از کمیته ارزیابی انتظار می رود، عبارت است از درک نظرهایی که کمیته متخصصان بر آن تأکید می کند.

ارتباط روشن، بین صداقت و مسئولیت پذیری، به دور از تداخل علایق، از اولین موارد با اهمیت است. کمیته ارزیابی، باید ضمن استقلال دانشمندان در حال داوری و مؤسسات آن ها، نمایندگی سیاست های مؤسسه ای که آن ها را انتخاب کرده اند، عهده دار گردد. نظام بودجه تحقیقات در لهستان، که در سال ۱۹۹۱ ایجاد شد، فاقد چنین ملزوماتی است. ۵۰٪ اعضای کمیته دولتی تحقیقات علمی (که هزینه های آن از بودجه تحقیقات تأمین می شود)، از دانشمندانی هستند که کمیته علمی آن ها را انتخاب کرده است.

### اعمال فشارهای روانی و اجتماعی

در نظام ارزیابی ایده آل، کمیته ارزیابی (دادگاه عادل)، آشکار از داوران (شاهدان) جدا شده است. از داوران انتظار می رود به بیان روشن عقاید خود بپردازند؛ عقایدی که بر دارا بودن اطلاعات کافی و کامل از دانشمندان که می خواهند مورد ارزیابی قرار دهند، متکی باشد. ممکن است عقاید آن ها درونی و ذهنی باشد، اما نباید تحت تأثیر هیچ گونه فشار خارجی قرار گیرد. یکی از پیش نیازهای این موضوع، آن است که داوران به تصمیم نهایی (قضات و حکم نهایی) کمیته ارزیابی نباید کاری داشته باشند و مسئولیتی در این زمینه پذیرند.

داوران یا اغلب تحت فشار هستند یا نمی توانند عقاید خود را به راحتی ابراز کنند. در دوره حکومت توتالیتار در لهستان و سایر

کشورهای اروپای شرقی، این اعمال فشارها، ابعاد سیاسی و ایدئولوژیک داشت و به‌ویژه درباره رشته‌های فلسفه و علوم انسانی اعمال می‌شد. پس از ایجاد تغییرات دموکراتیک در منطقه، این رسم برانداخته شد و فشارها حذف گردید، هرچند که برخی فشارهای ایدئولوژیک هنوز پابرجا هستند. جای تعجب است که اصلاحات سیاسی در ایالات متحده آمریکا، بر بررسی دقیق و سیاست علمی آن کشور تأثیری نداشته است. پس از حذف فشارهای دولتی، دیده می‌شود که بسیاری از انواع اعمال فشارها که در دوره نظارت بوروکراتیک مرکزی ایجاد شده‌اند، هنوز در تفکر افراد دانشگاهی و محققان وجود دارند. در پنجاه سال گذشته، سنت نقد علمی، کاملاً فراموش شده است. اعضای کمیته علمی در بیان انتقادی عقاید خود درباره کار دیگران تردید می‌کنند، حتی هنگامی که این عقاید کاملاً محکم و منطقی هستند. ارزیابی‌های غیرمنصفانه منفی‌نگر و از روی عناد، نادرند. از سوی دیگر بخش عمده‌ای از عقاید که براساس تعاریف، مثبت هم ارزیابی شده‌اند، به ترفیع‌ها، حمایت‌های مالی و سایر تصمیم‌های اساسی برای کسانی که در حال داوری شدن هستند، مربوط می‌شوند. درحالی که داور (ان) قلباً با این نوع ارزیابی موافق نیستند. داورانی که بررسی منفی دارند، اغلب به دلیل رفتار غیرانسانی کمیته علمی، منزوی و طرد می‌شوند.

یک دلیل برای این بررسی‌های غیرمنصفانه، وجود ارتباط ناصحیح بین داوران و کمیته‌های ارزیابی است، به این معنی که کمیته‌های ارزیابی، می‌کوشند با اتکا به نظر داوران، از خود سلب مسئولیت کنند. در لهستان، کمیته دولتی تحقیقات علمی، که بودجه تحقیقاتی را توزیع می‌کند، اغلب تصمیم‌های خود را براساس جمع‌بندی امتیازهای یک یا دو داور انجام می‌دهد، بدون این‌که بررسی‌های آن‌ها را تجزیه و تحلیل کند. بنابراین داوران در موضع قاضی قرار گرفته‌اند و به ابراز نظر صریح خود، تمایلی ندارند، زیرا می‌دانند با این کار، ممکن است دانشکده‌ای را از بودجه‌اش محروم کنند.

چنین اعمال فشارهایی به درجات بالاتری در ارزیابی اعطای مدارک یا ترفیع به مدارج بالاتر تأثیر می‌گذارند؛ حتی وقتی که موضوع رساله یا نتیجه پایان‌نامه در میان باشد، خطاهای آشکار یا ناکارایی آن مشخص می‌شود. اغلب داوران، با ایجاد کمیته‌های علمی مخالفند یا با آن دشمنی می‌ورزند. این موضوع، عملاً محفوظ نگاه داشتن نام بررسی‌کنندگان یا ارزیاب‌ها را در نظام خشک موجود و کمیته‌های کوچک علمی، مخدوش می‌کند.

تغییر این مفهوم کار دشواری است. تعریف دقیقی از برخورد عقاید و علائق و تعدیل فرایندهای ارزیابی، ممکن است مفید باشد، اما این کار هنگامی مؤثر خواهد بود که فرهنگ واقعی ارزیابی دقیق، در کمیته‌های علمی ریشه دار شود.

## نگاه پایانی: برخی ابهامات

در این جا به برخی ابهامات اشاره می‌شود:

۱- در نظر گرفتن فضای بُرداری  $X$  به مثابه فضای موفقیت و اعتبار ایده جالبی را القا می‌کند، اما نحوه مقایسه دو بُردار در فضای

بُرداری در این مقاله، چندان روشن نیست. بُردارها را می‌توان با هم جمع یا از هم کم کرد؛ همچنین برای بُردارها، ضرب‌های گوناگونی تعریف شده است، لیکن مقایسه دو بُردار، فقط براساس اندازه آن‌ها که کمیتی کلی است، صورت می‌گیرد. به این ترتیب می‌توان گفت،  $n$  طول بُردار  $x$  از طول بُردار  $y$  بیشتر یا از آن کمتر است، ولی رابطه در  $x < y$  در بُردارها معنی دار نیست. از طرفی، اگر بُرداری رابه مثابه تایی مرتبی همانند (با این  $x = (x_1, x_2, \dots, x_n)$  نظر بگیریم، آن‌گاه برای مقایسه دو بُردار نظریه  $x, y$  مشخصات، باید از تعاریف ترتیب در مجموعه‌ها استفاده کنیم که شاید با توجه به محتوای مقاله، ترتیب قاموسی برای این کار، از همه مناسب‌تر باشد، زیرا در این تایی ترتیب، عناصر در هر حال، مرتب، تک‌به‌تک با هم مقایسه می‌شوند. این ابهام در مقاله، کاملاً به چشم می‌خورد.

۲- عوامل وزنی و مقادیر آن‌ها، با ابهام عنوان شده‌اند. شاید در این‌گونه روش ارزیابی، یکی از مهم‌ترین اقدامات، یافتن مقادیر (وزنی)  $f$  باشد که در مقدار فراوانی هر موضوع (پ) ضرب می‌شود. چگونه دو مجله معتبر علمی با یکدیگر مقایسه شده‌اند و به هر یک عامل وزنی مشخصی داده شده است؟

۳- موضوع ارزیابی آثار تحقیقاتی، کار بسیار دشواری است و باید به شاخص‌های انسانی در آن توجه کرد. به همین دلیل، به روشی پویا (دینامیک) نیاز داریم که بتواند ضمن کمی کردن تعدادی از کیفیت‌ها، جریان تحول فرد در طول زمان را نیز مدنظر قرار دهد. اگر در هر سال، همه پژوهشگران ارزیابی شوند، در سال‌های بعد، باید به رشد فعالیت‌های آن‌ها در سال‌های گذشته نیز توجه شود. اگر دانشمندی مانند «اینشتین»، وقت زیادی صرف ریاضی کردن ایده نسبت عام می‌کند و در این سال‌ها تولیدات علمی محدودی به جامعه علمی عرضه می‌نماید، این زمان، باید در محاسبه کیفیت و حتی کمیت کار او ملحوظ شود، زیرا او در این مدت، سرگرم تولید مفاهیم تازه‌ای بوده است که یا هنوز کاملاً پایان‌نیافته‌اند یا از انتشار ایده‌های اولیه آن، صرف‌نظر شده است. به‌طورکلی می‌توان گفت، تحقیقات به مثابه سیستم، بیشتر به یک سیستم دینامیکی پویا شباهت دارند (زیرا در طول زمان، دچار تحول و تغییر می‌شود) و نمی‌توان آن‌ها را با روش‌های ایستا ارزیابی کرد.

۴- جالب‌ترین بخش‌های مقاله از قسمت‌هایی تشکیل شده که مؤلف به نحوه داوری و ارزیابی، محدودیت‌ها، فشارهای روانی و اجتماعی و مواردی از این نوع پرداخته و توصیه‌های روشن و دقیقی در این باره عرضه کرده، گرچه در انتها، فقط به ذکر عبارت‌هایی چون: «باید در روش‌های ارزیابی تجدیدنظر شود»، «با افزایش تعداد داوران، می‌توان به قسمتی از مشکلات داوری غلبه کرد» و «داوران باید از کمیته‌های ارزیابی مستقل باشند»، اکتفا نموده است، حال آن‌که موضوع داوری و ارزیابی، هنوز هم یکی از موضوعات پرابهام، در سطح جهان محسوب می‌شود. چگونه می‌توان در عملی که انسان در آن نقش دارد، برخی عوامل انسانی نظیر ارجحیت دادن به دوستان و نزدیکان، پیش‌داوری، ملاحظات شغلی و... را که فقط از انسان صادر می‌شوند، حذف کرد؟ دیدگاه سیستمی و سینرژیتیک اعتقاد دارد که

سیستم‌ها از قابلیت خودسازمان‌دهی برخوردارند (البته نه همه سیستم‌ها، بلکه آن‌هایی که غیرخطی هستند). تلاش عمده سیستم‌های موجود که به هیچ‌وجه با آن‌ها به صورت پویا برخورد نمی‌شود، این است که با صدور دستورالعمل‌ها، آیین‌نامه‌ها، بخشنامه‌ها و قوانین خدشه‌ناپذیر؛ به اصولی دست یابند که همه چیز در آن تعریف شده، دقیق و بدون ابهام باشد. لیکن تا امروز، حتی یک نظام اداری موفق برای انجام دادن کارهای معمول ایجاد نشده است. اکنون در نظر بگیرید که موضوع ارزیابی و داوری، با چه دشواری‌هایی روبه‌رو است. محکم‌ترین قوانین، قابلیت نفوذ دارند و امکان تغییر جهت آن‌ها در مسیری خاص و تعبیر و تفسیر آن‌ها به روش‌های دلخواه وجود دارد. این مشکل، معمولاً با ذکر این جمله «تعبیر و تفسیر این آیین‌نامه در حوزه مسئولیت ... است» که به جای چند نقطه در این عبارت، می‌توان فرد یا سازمان معینی را عنوان کرد، برطرف می‌شود. لیکن این اقدام نیز تاکنون پاسخ‌گویی مشکل نبوده است، زیرا در این‌گونه اقدامات، از پایه متناقض هستند و وجود تناقض‌هایی از این نوع، در اساس نظام‌های نامعطف و ایستا نهفته است. در بسیاری از کارها، فقط وجود چند قانون عام لازم است، ولی اصلاح مسیر و نحوه انطباق آن با شرایط جدید، باید تا حد زیادی در اختیار خود آن نظام قرار گیرد و این چیزی است که نظام‌های نامعطف فعلی، توان پذیرش آن را ندارند. در هر حال، موضوع حفظ بی‌طرفی در داوری ارزیابی نیز، مقوله‌ای از همین نوع است و برخورد نامعطف با این پدیده انسانی، کارساز نخواهد بود. **د** عقیده بر این است که علوم انسانی را در قالب الگوهای ریاضی نمی‌توان گنجانند و کسانی که از این سخن دفاع می‌کنند، بیشتر در شاخه علوم انسانی تخصص دارند و دخالت ریاضیات در این رده ارزشمند از علوم و معارف بشری راه نوعی کم‌توجهی یا بی‌توجهی به قوانین عام متکی بر انسان تلقی می‌کنند. هرچند که این نوع تلقی، چندان هم دور از فعالیت نیست، اما باید توجه داشت که جوامع انسانی نیز، خواسته یا ناخواسته، از قوانین عامی پیروی می‌کنند که برخی از اصول آن را جامعه‌شناسان و روان‌شناسان شناخته و بررسی کرده‌اند و هر جا که قوانین عام حاکم باشند، ریاضیات نیز قدم به میدان می‌گذارد. بنابراین برای شناسایی نظام‌های بسیار پیچیده‌ای چون انسان و جوامع انسانی، باید از هر نوع ابزار ریاضی‌سود برد و ریاضیات کلید ارزشمندی برای این کار است. نظام ارزیابی نیز به دلیل دخالت انسان در آن، باید تا حد امکان شناسایی شود و نقاط گرهی، رفتارهای آشفته (chaotic) و تمایلات و جهت‌گیری‌های آن، باید شناسایی شوند و مورد بررسی قرار گیرند. این مقاله نیز، فقط به همین منظور عرضه شده است و به‌رغم ابهامات و ضعف‌های موجود در آن، می‌تواند به مثابه ایده و نظری در این باره، مورد توجه قرار گیرد.

**پیوست**

**تعریف فضا برداری**

فرض کنیم  $F$  میدان و  $V$ ، گروه آبدلی جمعی باشد. فرض کنیم برای

ضرب هر عضو  $x$  از  $F$  در هر عضو  $v$  از  $V$ ، یک ضرب عددی تعریف شده باشد، به طوری که عضو منحصر به فرد  $xv$  از  $V$  حاصل شود، مشروط به آن‌که به ازای هر  $x$  و  $y$  از  $F$  و هر  $u$  و  $v$  از  $V$  داشته باشیم:

$$x(u+v) = xu + xv \quad (1)$$

$$(x+y)u = xu + yu \quad (2)$$

$$(xy)u = x(yu) \quad (3)$$

$$1u = u \quad (4)$$

که در آن 1 عبارت از است عضو یگانه میدان  $F$ ، در این صورت،  $V$  یک فضای برداری روی میدان  $F$  نامیده می‌شود که گاهی آن را زوج  $(V, F)$  نمایش می‌دهند.

**تعریف فضای متریک (متری)**

فضای متری مجموعه‌ای است مانند  $X$  و یک تابع  $d: X \times X \rightarrow R$  که در شرایط زیر صدق می‌کند:

$$d(x,y) > 0$$

$$d(x,y) = 0 \Leftrightarrow x=y$$

$$d(x,y) = d(y,x)$$

$$d(x,z) < d(y,x) + d(y,z) \quad \text{نامساوی مثلثی}$$

تابع  $d$  را متر فضای متری گویند.  $d(x,y)$  نیز، به‌عنوان فاصله بین  $y$  و  $x$  تعبیر می‌گردد. توجه شود که در مجموعه حاصل ضربی  $X \times X$  همه عناصر به صورت زوج‌های مرتب  $(x,y)$  هستند که در آن  $x \in X$  و  $y \in X$  و

**منابع و مأخذ:**

- Frankel, M.s., Cave J. (eds.), Evaluationg Science & Scientists, Central European University Press, Budapest 1997.
- Dickinson, J.,P., Science & Scientific Researchers in Modern Society, Unesco, Second edition, 1989.
- Michaelis, A.R. & Harvey, H. (eds.) Scientists in Search of their Conscience, Springer-Verlag.