

批判性思维定义的评价标准及其应用

梁润成

摘要: 学者们提出了众多关于批判性思维的定义, 而选择采用哪一个定义是一个富有争议的问题。约翰逊和汉比也注意到了这一点, 他们另辟蹊径, 从元视角出发审视批判性思维的定义, 尝试提出一套用于评价定义的标准。本文同样采取这一元理论视角, 但试图在约翰逊和汉比提出的标准基础上加以完善, 并将其应用于批判性思维定义的评价之中。具体而言, 他们的标准以单个定义项为评价对象, 忽视了定义提出者对定义项的阐释文段, 没有充分重视对定义文段集合中“思维”论述的考察, 也未提出用于比较不同定义文段集合的标准。为此, 本文从“批判性”和“思维”这两个词在英语中的词源出发, 提出了界定批判性思维时所面临的三个元问题, 并依据定义文段集合对这些元问题回答的理论可能性, 提出了内部融贯标准与对比标准。最后, 本文将新标准应用于杜威、法乔内与拜林的定义文段集合, 结果发现三者均符合内部融贯标准, 而在对比中可以看出, 杜威的文段在回应“思维”及“思维类型比较”的问题上采用显性回答, 而在“批判性”的阐释上则更有“厚度”, 因此在实践指引力上更具优势。

关键词: 思维; 批判性; 批判性思维的定义; 评价标准; 杜威

中图分类号: B81

文献标识码: A

1 引言

我们拥有众多关于批判性思维的定义, 这些定义并不相同。要相信哪一个批判性思维的定义, 是一个颇具争议的问题。本文旨在提出一套评价标准, 并对定义进行评估。

这项工作面临两种可能削弱其研究意义的观点: 首先, 学者们基于各自的理论视角提出不同的定义, 每个定义都有其独特价值。因此, 定义的多样性并非问题, 甚至应鼓励“百花齐放”。这一观点在一定程度上是合理的, 因为学术研究确实应包容多元。然而, 探讨评价标准的做法并不与此相矛盾。评价标准本身并不预设任何特定定义, 而是站在这些定义之外, 以更宏观的视角审视整体。因此, 研究定义的评价标准实际上体现了一种独特的学术视角。

收稿日期: 2024-11-20

作者信息: 梁润成 广东第二师范学院马克思主义学院
Garyliang924@gmail.com

其次,在批判性思维领域,法乔内(P. Facione)主导的《德尔斐报告》(*Delphi Report*)([13])与“百花齐放”的理念似乎存在一定张力。该报告汇集了众多专家,在批判性思维的定义上达成共识,这一成果无疑具有里程碑式的理论价值。但问题在于,这种共识性定义是否意味着不再需要进一步探讨批判性思维的定义?答案是否定的。原因在于,《德尔斐报告》在提出其定义时,并未充分讨论或引用其他学者的相关研究。我们不禁要问:其定义与前人的有哪些异同?优势又体现在哪些方面?仅从报告本身,很难获得这些问题的答案。约翰逊(R. Johnson)和汉比(B. Hamby)也提出了类似质疑,他们担忧《德尔斐报告》的定义未能充分与其他替代定义建立关联。([20],第420页)因此,虽然《德尔斐报告》可视为“百花”中的一朵,但它的出现并未终结批判性思维定义的研究,反而激发了我们对该概念进行更深入的探讨。

在回应反面观点之后,本研究具有以下正面意义:第一,评价标准为分析众多定义提供了一个框架,使我们能够系统梳理这些定义,并识别其差异、相似性及潜在冲突;第二,它为判断定义的质量提供了尺度,帮助我们在不同定义之间进行对比和评价;第三,只有当定义得到充分辩护后,与之相关的实践活动才能获得理论上的支撑。

接下来,本文首先对约翰逊和汉比的工作进行分析与评述;第三部分提出一个新的评价标准;第四部分将该标准应用于三个具有代表性的定义;最后一部分对全文进行总结。

2 约翰逊和汉比的标准

2.1 约翰逊和汉比标准的阐发

“元层次方法”(meta-level approach)是针对当前批判性思维定义混乱、缺乏统一标准的问题而提出的一种规范化的定义策略。元层次方法的四个步骤:新定义的提出者首先要指出现有定义的缺陷或不足,以说明提出新定义的必要性;随后,应明确自己所提出的定义属于哪一类型,并列出该类型定义应满足的具体标准;提出者正式给出自己的定义;最后需论证其定义符合前述标准。([20],第418页)

约翰逊和汉比在其论文中把批判性思维界定为“纲领性定义”(programmatic definition)。他们借用了谢夫勒(I. Scheffler)的相关概念,以说明“批判性思维”的定义应具有的功能定位与理论指向。此类定义不是词典式定义(lexical definition),也不是约定性定义(stipulative definition),而是提出一种新的、规范性的理解方式,引导人们如何思考和实践。([25],第22页)

在上面这些理论铺垫下,约翰逊和汉比为评估批判性思维定义提出了三个标准:

第一,不得循环,即定义项不能包含被定义项或其同义词;

第二，范围恰当，既不过宽到把通常不算批判性思维的活动也囊括，也不过窄到把公认实例排除在外；

第三，必须明确“批判性”（critical）这一术语的界定。

（[20]，第422-423页）

他们并未依据这三条标准提出自己的新定义，而是把标准作为“元层次”工具，用来检验和质疑现有定义。基于这个理论框架，他们认为当前对批判性思维的定义普遍存在四种不足：首先，许多定义未能阐明“批判性”一词的核心含义；其次，存在概念混淆，将批判性思维与问题解决、决策、元认知或创造性思维等其他认知活动混为一谈，忽视了语言中不同术语所承载的重要区分；第三，一些定义过度泛化，甚至将批判性思维理解为一切好思维的代名词，使其失去理论特异性和教育可操作性；最后，多数定义片面聚焦于信念的获取（如“决定相信什么”），却忽略了同样关键信念维护与修正过程，尤其是通过与他人互动来检验和调整自身观点的能力。这些缺陷共同导致现有定义缺乏清晰性、区分度和实践指导力。

2.2 对约翰逊和汉比标准的评论

首先，约翰逊和汉比的评价标准虽然突出了“批判性”的核心地位，但在一定程度上忽视了“思维”作为本体论基础的重要性。尽管他们主张定义中的每个词项都应对整体意义有所贡献，但在实际操作中，“思维”这一概念主体并未得到充分体现。“批判性”并不是独立存在的行为，而是对“思维”的一种限定，用以界定特定类型的思维方式。因此，如果评价仅聚焦于“批判性”这一修饰，而忽略“思维”这一承载主体，就可能导致对概念必要条件的理解出现结构性偏差。

其次，该标准主要依赖对单句定义的分析，缺乏对语境的充分关注。实际上，定义的意义往往是在解释、说明和例证的配合下共同生成的，而非由单一句子独立决定。脱离语境进行评估，容易造成误解。约翰逊和汉比在附录中以单句形式呈现大量定义，恰恰反映了其方法的局限。虽然这种方式适用于外延清晰的词典式定义，但对于“批判性思维”这类具有规范性和纲领性的概念而言，则难以全面把握其内涵。

最后，该标准在区分和甄别不同定义优劣方面也显得不足。仅仅强调“批判性”一词的出现，并不能有效区分理论深度不同的定义，因为学界对这一概念的理解本就多样。评价定义不应停留在要素拆解层面，而应通过多角度的比较与分析来进行。正如选择衣物需要考虑多种因素，界定概念也需要在不同解释之间加以权衡，从而识别出那些在解释力上更具优势的定义。

3 新标准

3.1 新标准的阐发

约翰逊和汉比指出,“批判性”一词具有三种主要含义。([20],第424页)第一种出现在日常语境中,意指“过分负面”或“挑剔”,他们明确排除了这一含义在批判性思维中的适用性。第二种含义是“危急的”或“至关重要的”,如“病人情况危急”或“关键阶段”;这一含义虽强调重要性,但并未触及批判性思维的核心内涵。第三种才是他们认为真正适用于批判性思维的含义,即“展现出良好的判断力”,其具体表现包括能够权衡证据、分析各方观点的利弊、并最终形成有根据的判断。

约翰逊为他所提出的“批判性”一词的第三种含义提供了词源学依据。“critical”源自希腊语“kritikós”,意指“能够判断、辨别或作出决定”,而“kritikós”又来自动词“krínein”,其本义为“分开、区分、判断、决定”。([19],第49页)在古希腊,该词常用于形容那些具备辨别力、能够进行理性判断的人,尤其用于哲学、文学和法律等领域。我认为约翰逊的这一观点具有较强的说服力。

然而,约翰逊的研究尚显不足,因为他忽视了对动名词“thinking”的词源学剖析。追溯至古英语,“think”的词根实则源自两个语义迥异却紧密相关的动词:“þencan”与“þyncan”。前者构成了现代“think”的直接前身,代表“主观能动”维度,指涉主体在脑海中的主动操作(如构思、判断或打算),即“我在头脑中生成观念”(如 I þohte……);后者则代表“客观显现”维度,此时主语并非思考者而是被感知的客体,表达某事在人心“显得”如何,是一种被动唤起的印象(如 It þuhte mē……,意为“某事在我心中呈现出某种样子”)。尽管这两组词根在中古英语时期合并为“thinken”,但“þyncan”所含的“似乎/显得”之意逐渐式微,仅残存于“methinks”等古式表达中,而“þencan”的“思考”义项则占据了主导。([5, 6, 16, 22])现代英语中的“think”虽直接承袭自中古英语的“thinken”,但这一演变过程深刻揭示了该词内在的双重维度:它既包含主体的主动操作,也蕴含客体的被动显现。

这里对“think”在词源上的分析,为我们提供了一个理论背景去审视“critical”所体现的判断活动,“critical”需要在“think”框架下去理解,即作出有依据判断这种行为是在主动操作以及被动显现的双重背景下进行的。这为我们分析定义文段集合提供了三个元问题:

第一,如何从主体的主动操作与对象的被动显现两个维度来理解思维活动?

第二,作出有根据的判断与其他类型的思维活动相比,其根本差异体现在哪里?

第三,作出有根据的判断在主体主动操作与对象被动显现这两个方面分别如

何表现？

这三个问题从“思维”“思维类型的比较”以及“批判性”三个维度来评价定义文段的整体表现，而这三个维度共同构成了理解批判性思维的语义框架。元问题在某种程度上继承了约翰逊和汉比的标准，上文指出他们要求划清批判性思维概念的边界，这对应于元问题二；应注重对“批判性”论述的评价，这对应于元问题三。我增加的是对“思维”部分论述的评价，并指出“批判性”是修饰的，不应作为一种独立的行动，须置于“思维”框架内方可理解，这对应元问题一。基于定义文段集合对三个元问题的回答，我提出相应的标准来进行评价。元问题是标准背后的原则。该标准通过考察定义文段集合回答元问题的方式，对文段进行类型划分，并将某些回答特征标记为正面，另一些标记为负面。由此，我给出如下标准¹：

一个定义文段集合是内部融贯的，当且仅当它对上述三个问题作出了显性或隐性的回答，且不存在循环或前后不一致的情况。

上述标准针对某一定义文段集合，我还提出以下对比标准去衡量不同的文段，

规则一：对于两个定义文段集合，若其中一个未对上述问题作出显性回答，而另一个作显性回答。在比较时，显性文段优于隐性文段。

规则二：若两个定义文段集合就某个问题的回答均为显性，则比较它们在回答上的“厚度”，回答“更厚”的文段更优。

第一，什么是“定义文段集合”？元问题是新标准所设立的分析工具，旨在帮助我们评价文段中的相关论述。所谓“定义文段”，是指文本中在语义上回应某一元问题的语句，它可以是一句话，也可以是一段文字。这类文段可能出现在同一部著作的不同章节中，也可能分散于作者的多部著作中，或体现于其他学者对其理论的阐释性工作中。需要强调的是，学者在写作时未必有意识地回应全部三个元问题，因此某部著作中可能仅包含回应其中某一元问题的文段，而缺少对其他元问题的回应。“定义文段集合”是由这些回应元问题的文段构成的整体，并通过元问题所建立的语义框架将其加以统摄与整合。

第二，内部融贯标准有何种功能？三个元问题构成了对批判性思维定义文段集合的最低语义要求，它们规定了文段内部应具备的基本语义单元。“不矛盾”和“不循环”则属于所有文段写作的共性原则，一旦违背这两项原则，文段便会变得不可理解。因此，该标准并非用于评判定义文段集合内容的优劣，而是用于确保其具备可理解性的基本条件。只有当一个文段具备可理解性时，它才可能与其他文段进行比较。

¹感谢匿名审稿人就这部分内容提出了十分有益的修改意见，获益良多。

第三,什么是“显性回答”?它指在定义文段集合中,作者明确地、直接地讨论某个问题,并试图给出清晰的答案——而不是仅仅在谈论其他概念时顺带提及或间接暗示。在实际文本中,显性回答通常有三个特征:设有专门的章节或段落,直接聚焦于该问题;使用明确的关键词(比如“判断”“思维”等)来标识讨论主题;对这些关键词进行解释或界定,而不是仅仅使用它们而不加说明。支持这种做法的理由可以在格赖斯(H. P. Grice)的理论中找到。日常对话之所以能够顺利进行,是因为交谈双方默认遵守“合作原则”及其四个准则(数量、质量、关联、方式)。(〔15〕)在某些强调信息明确传递的语境中(如学术写作),作者在文本中对某一问题进行“显性回答”时,正是在遵守方式准则:通过设立专门段落、使用明确关键词、对概念进行界定,使表达清晰、有序、无歧义。这种结构化的回应方式最大限度地降低了读者对作者立场进行过度推测或误读的可能性,使信息传递更为高效。

第四,什么是“厚”的定义文段集合?威廉斯(B. Williams)尝试界定“更厚”(thicker)的概念。(〔26〕)他指出,像“勇敢”“残忍”这样的厚概念,不只是表达价值判断,还包含具体事实描述——比如“他用脏话打断老奶奶发言”,这句话既说明了发生什么,又暗含道德批评;而像“好”“坏”这样的薄概念,则只给出抽象评价,比如“这是不对的”,却不告诉你具体错在哪里。除了“厚概念”外,我们还有“厚描述”,赖尔(G. Ryle)把“厚描述”(thick description)刻画为:在描述人类活动时,不只记录身体动作,还纳入意图与目的等细节,从而帮助我们理解这些活动。(〔24〕)在格尔茨(C. Geertz)的术语里,“厚”意味着这些细节具有文化特定的意义。(〔14〕)格尔茨借用了赖尔的“厚描述”概念,但将其从分析个体行为的方法扩展为理解整体文化的路径。赖尔关注的是个体行动背后意图的层次差异,例如区分“眨眼”与“抽搐”时所涉及的意义结构;而格尔茨则把重点转向文化事件中多重意义视角的交织,例如在摩洛哥部落抢羊的情境中,犹太人、柏柏尔人以及法国殖民者等不同群体各自赋予事件的不同理解。简而言之,赖尔用“厚描述”回答的是“一个人在想什么”,而格尔茨则用它来揭示“人们共同在做什么,以及这一行动在文化上意味着什么”。可见,我们用“厚”来修饰某个对象时,其表明这个的内容更丰富,细节更多。在此,我想把“厚”用来修饰定义文段集合。类似地,定义文段集合的目标在于界定一个概念。“薄”文段仅给出对元问题的回答,而不对其内部要素进行进一步说明;相反,“厚”文段集合不仅回应元问题,还会对答案中涉及的要素展开阐释,例如解释关键词、提供例子、给出理由、描述形成过程或进行概念对比等。因此,“厚”文段集合的特点在于,它在回答元问题时采用了更为丰富的修辞结构,以增强概念的清晰度和内容的丰富性。

第五,如何判断哪个文段集合更“厚”?在界定“厚”文段集合时,我依据其

在修辞上的表现力进行判断,但这些表现力也需聚焦在某个价值目标上,不然只是无意义的炫技,所以在比较文段集合“厚度”时,需引入修辞结构之外的衡量标准,即考察文段的实践指引力。所谓“实践指引力”,是指定义文段集合在理论上能够涵盖的实践情境范围,以及其对这些情境所提供的规范性指导程度。实践指引力越强,其适用的情境越广,所提供的规范指引就越层次分明、结构就越丰富。之所以选择这一尺度,是因为批判性思维的终极价值取决于其理论能否在现实情境中帮助个体更有效地处理问题和作出判断。正如法乔内所指出的:“教会人们做出明智的决策,就等于赋予他们改善未来的能力,使其成为社会的贡献者,而非负担。”([12],第1-2页)

3.2 对可能反驳的回应

反驳一:新标准强调必须对“思维”本身进行论述,否则定义不完整。然而,“思维”作为抽象且具有争议的概念,若要求每个批判性思维定义都先界定“思维”,则可能陷入无限回溯:为定义“批判性思维”,需先定义“思维”;为定义“思维”,又需定义“意识”或“心智”等更基础的概念。这不仅不切实际,也可能使定义变得臃肿而失去实用性。

对反驳一的回应:要求对“思维”进行说明,并非要求重新构建一套完整的“思维理论”,更非陷入概念冗余。所谓“冗余”是指无信息增量的重复,而此处的要求恰恰是为了避免信息缺失。因为“批判性”是对“思维”的限定,若连“思维”在此语境中指涉何种活动都未澄清,则“批判性”就失去了所依附的载体。在定义文段集合中,应至少要说明“思维”的工作性定义,列举不同的“思维”类型,为从“思维”过渡到“批判性”提供说明。新标准所要求的,正是这种语义上的基本交代,以确保定义在论述链条上是融贯的,而非语义悬空的口号。

反驳二:新标准的元问题要求批判性思维的界定应服务于理论阐发目标。但实际上,定义具有多重功能——有的用于课程设计,有的用于心理测量,有的用于哲学探讨。一个适用于标准化测试的定义与一个用于哲学的定义在结构和重点上必然不同。用这两个标准去衡量功能迥异的定义,可能造成“错位评价”,忽视定义的情境适切性。

对反驳二的回应:新标准并不否认“批判性思维”的定义会因使用目的不同而有所侧重。然而,无论其功能如何,只要使用“批判性思维”这一术语,就必须在语义上回应三个元问题。这些问题构成了“批判性思维”作为可识别、可交流概念的最低语义门槛,也正是内部融贯标准所要求的。无论处于何种语境,对该概念的澄清都是前提条件。此外,新标准并未忽视语用与实践维度的评价。显性与隐性文段的区分,主要体现语用层面的差异:它影响读者对文段内容的理解与把握。而文段“厚度”的比较,则侧重实践指引的维度,即判断文段在理论上

对实际应用的指导力。由此可见，新标准并非割裂语义、语用与实践，而是以语义为基础，同时整合语用和实践特征进行整体评价。

反驳三：新标准强调对定义文段集合的整体理解。然而，这些判断本质上高度依赖评价者自身的理论立场、学科背景、语言敏感度甚至价值预设，因此，评价结果很可能因人而异。

对反驳三的回应：内部融贯标准及对比标准规则一的应用主要基于元问题来检测文段。文段作为已完成的产物，具有客观性。是否对某些问题作出了显性或隐性回答，可以完全依据文段中的词句找到答案，主观解读空间有限。至于对比标准的规则二的应用，虽然可能存在主观解读的差异，但新标准通过限定评价范围、明确提取规则并要求依据文本证据，将解释约束在文段之内，从而引导评价聚焦并减少任意性。新标准的目的不是消除所有解释差异，而是确保评价立足于文段本身，并可被他人复核与质疑。

反驳四：定义文段集合的选择极易影响新标准的分析评价结果；若所选文段存在遗漏或数量不足，可能导致分析结果出现偏差。

对反驳四的回应：该反驳指出了新标准对所选文段的依赖性，其所警示的风险确实可能存在。为此，本文在文段选取上力求客观呈现作者的核心观点，主要依据能够体现其主要立场的代表性原著及权威的二手研究文献，以尽可能减少因文本选择范围有限而对分析评价结果造成的偏差。

4 新标准的应用

4.1 定义文段集合简述

我选择了三个定义项，并分析了它们背后的定义文段集合。这些定义项在语篇结构中起到统领和引导作用，为定义文段集合的展开提供了方向。

拜林 (S. Bailin) 和巴特斯比 (M. Battersby) 认为，批判性思维是：

为了获得合理判断，对议题进行的仔细审查。 ([1], 第 4 页)

他们将批判性思维置于论辩领域，认为其功能在于考察议题背后的多方论证，评估其优劣并作出合理判断。他们从论辩及认识论的视角，理解批判性思维中的规范。同时，他们不希望使用“技能”来界定批判性思维，因为该词带有显著心理学特征，其水平可高可低，不完全符合批判性思维的要求。只有当技能达到高水平并符合规范性时，才可在批判性思维的论述中使用“技能”一词。([2]) 他们希望批判性思维成为一个规范性的理论框架，以指导教育实践。²因此，其没有使用技能与倾向来概念化批判性思维，而是依赖背景知识、操作知识、关键概念、启

²感谢拜林教授在 ECA 2025 茶歇期间的详细讲解，让我对其理论有更深入的认识。

示法（包括策略和程序）以及心智习惯来表达批判性思维，从而避免将其与人的心理状态直接关联。（[3]）

法乔内将批判性思维定义为：

一种有目的的、自我调控的判断过程³，这一过程表现为对信息的解释、分析、评估与推理，并对该判断所依据的证据性、概念性、方法论、标准性或情境性因素进行说明。（[13]，第3页）

正是基于这一界定，法乔内将批判性思维所涉及的技能区分为解释、分析、评估、推理、说明以及自我调控。除了技能之外，他还提出了一系列与这些技能相关的心理倾向。需要注意的是，法乔内旨在提出一个理想的教育目标，因此这里的技能应理解为达到“熟练”水平时能够产生规范结果的能力。

杜威（J. Dewey）最开始使用的词是“反思性思维”（reflective thinking）（[9]，第2页），这被通常理解为就是现今的批判性思维。他把反思性思维界定为：

对任何信念或假定的知识形式，根据支持它的根基和它所倾向的进一步结论，进行积极、持久和仔细的考虑。（[10]，第9页）

在1938年的著作中，他使用了“探究”（inquiry）一词，去取代反思性思维，并把探究作如下定义：

一种受控制或有方向的转化过程，它把一个不确定的情境转变为一个在构成要素及其关系上足够确定的情境，以致原来情境中的各个要素被重新整合为一个统一的整体。（[11]，第117页）

反思性思维与探究在核心过程上并未改变，但在定义的表达方式上有所差异。反思性思维侧重个人的内在精神操作，这一考虑信念的过程可能无需实际行动；而探究则在精神操作的基础上，强调对情境的改变，这种改变是实际发生的并带有规范意义。

之所以选择这些学者的定义文段集合，是因为他们具有代表性。杜威是第一位系统性提出批判性思维理论的学者（[17]），其工作代表了该理论的开端。法乔内的研究反映了当时学术界的共识，该报告的参与者包括恩尼斯（R. Ennis）、保罗（R. Paul）、李普曼（M. Lipman）及约翰逊等（[13]，第35页），他们均是批判性思维领域的权威人物。拜林和巴特斯比的观点则代表了当前学界对批判性思维的最新理解。

三个所选的定义文段集合满足内部融贯标准，主要理由在于，在搜集过程中未发现明显的矛盾或循环。当然，这种做法并非严格证明，不能据“未发现”即断

³在汉语中，judgment可译为“判断”“评判”或“判决”等，具体取决于语境：如“判决”偏法律语境，而“判断”较为中性。在该批判性思维定义中，judgment并非指静态结论，而是一种具有目的性与自我调控特征的动态认知活动。因此，将其译为“判断过程”更贴合其理论内涵。

言“必不存在”。不过，内部融贯标准本身并不要求形式化证明，而是立足于读者理解文本的角度，强调文本应满足的基本可理解性要求。因此，未发现显著矛盾或循环，已可在一定程度上支持这些文段符合该标准。进一步说，若所搜集的观点本身存在明显矛盾或循环，其理解应当受到阻碍；而从实际提取结果来看，不同文段的观点均具有基本的可理解性，这也为其满足内部融贯标准提供了佐证。

4.2 定义文段集合的对比：“思维”

在拜林和巴特斯比著作中，“thinking”绝大部分是与“critical”同时出现的。但也有以下几种情况，“thinking”不与“critical”同时出现：第一，“thinking”独自出现时，其作为一个动词名词去表达考虑某些对象，例如“thinking about argumentation”；（[4]，第 80 页）第二，他们在探讨批判性思维与认知谬误时，还引用卡尼曼（D. Kahneman）的“快思维”（fast thinking）和“慢思维”（slow thinking）。（[4]，第 181 页）他们并未设置专门的章节讨论“thinking”，而只是直接使用这一词语。这一点在他们 2015 年的著作中亦可得到印证——在书末按字母顺序列出的关键词表中，“T”项下并未出现“thinking”一词。（[1]，第 340 页）

法乔内在他的著作中也有三种关于“thinking”的用法，第一种是与“critical”并列，第二种是与其它修饰语并列“high-order thinking”（[13]，第 13 页），第三种是表明对象“thinking about an important or complex matter”（[13]，第 18 页）。但类似地，其文本没有专门部分解释“thinking”，只有这种直接使用。由此可推知，拜林与法乔内将思维理解为一种内在的主动操作：该操作具有特定的对象，遵循一定的方式，并产生相应的结果。

杜威对“thinking”的论述则更系统和全面。要先从他对反射弧理论的批判说起。他批判了传统心理学将刺激与反应割裂开来的机械论模型，认为传统反射弧理论将心理活动理解为一系列线性、分离的因果链条，忽视了有机体与环境之间的连续互动过程。（[8]）

设想一个孩童目睹燃烧的蜡烛并伸手触摸的场景。依据传统的反射弧理论，这一过程被机械地切割为三个离散阶段：刺激（视觉捕捉到火焰）、中介（神经冲动传导至大脑）以及反应（伸手触摸随即缩回）。在此模型中，刺激与反应被视为彼此割裂、线性递进的独立事件，仿佛其间不存在连续的互动或情境依赖。然而，杜威批判这种划分是人为且抽象的。他指出，“看到火焰”这一所谓的刺激，唯有置于正在进行的行动情境中方具意义。孩童并非被动地接收视觉信号，而是在探索新奇事物的主动活动中感知火焰。当触碰导致灼痛并引发缩手动作时，这一新经验即刻重构了他对火焰的知觉：火焰不再仅仅是有趣的亮光，而转化为“危险的源头”。

在杜威的理论中，“经验”（experience）既非单纯的主动行为，也非被动的接

受,而是有机体与环境之间持续互动的动态过程。经验表现为“行动”与“遭遇”的统一:人通过行动作用于环境,同时又受到环境变化的反作用;正是在这种双向互动中,人不断更新对环境的理解,并由此激发新的行动。除了交互性,经验还具有连续性——当下的经验既承继自过往,又并非一成不变。人在当下情境中能够主动反思并修正以往的经验,从而使经验成为一个在继承中不断调整、发展的整体。正如里加(F. Riga)所指出的,在传统观念中,经验是被发现能运用于实践的⁴记忆积累,它涉及主动行为或被动遭受,但在杜威看来,经验不只是过往的积累,更是在人与环境的交互过程中,蕴含着获取新知识可能性的行动。([23],第232页)

基于这一理论框架,杜威给出了“thinking”的定义:基于心智形成的心理状态序列(procession of mental states through his mind)。([10],第4页)心智并非某种固定的实体,而是一组动态的活动过程。语言本身也揭示了“心智”概念的多重意涵:它可以表现为记忆(“我想起某事”),体现为关注(“我惦记着她”),也可以指向意图(“我心中有目标”)。([18])可见,心智是思维的基础层,它为思维提供主动的操作,使人能够抓取和处理对象,进而采取行动。然而,思维并非完全主动的活动,其中还包含被动呈现的层次。譬如在“我想起某事”这一表述中,“我想”体现了主动操作,而“某事”的观念浮现则是被动的呈现。正是这种主动与被动的交织,构成了心理状态的连续序列。

由此,杜威区分了四种思维。([10],第4-8页)第一是“意识流”,如白日梦与幻想,其特点是想法自发、无序地在意识中流动,常在日常活动中不期而至。这类思维可理解为对主动操作的隐匿,但并未完全消失,仍保留着潜在的操作与选择能力。第二是“想象”的思维,指向未被直接感知的对象。例如,在讲故事时,叙述者与听者分别构思与重建连贯的情境结构,体现出思维的主动选择与组织功能。第三是“相信”的思维,在此过程中,个体决定何者值得信赖,从而形成既包含智力承诺又具有实践指向的信念;⁴这些信念指向外在对象,其合理性需在事实、原理或规律层面上接受检验。第四是“反思性思维”,它是“相信”思维的一种特殊形态,具有三项特征:其一,思维以连续关联的判断展开;其二,该判断链条指向明确的结论;其三,结论依赖探究而非臆测。杜威以哥伦布反对“地平论”为例加以说明:后者源于习惯与表象,而哥伦布则基于“地球为圆”的假设并通过航行加以检验,尽管结果超出其原初目的,但恰恰体现了反思性思维以探究为核心的本质。

⁴这指一种知行合一的状态:个体既在理性上审慎认同某一命题,又在实践中以此为指导;但这并不意味着该命题本身必然具备客观真理性。

4.3 定义文段集合的对比：“批判性”

在追溯“批判性”的英语词源时，我援引约翰逊的研究，将其核心内涵界定为“作出有根据的判断”。必须澄清的是，“批判性”作为一种修饰性特质，虽体现了思维中一种特殊的主动运作，但并非所有的主动思维活动都等同于判断。这种主动运作具有显著的对象依赖性：它并非凭空产生，而是仅在特定对象显现时被激活。换言之，批判性既蕴含于主动操作的特性之中，又依赖于被动呈现的对象。基于此，本文将从“判断的对象”与“判断的手段”两个维度，系统梳理所选定义文本中关于“批判性”的论述。

在拜林和巴特斯比的理论中，他们将“议题”（issue）作为判断的对象，定义为“一种挑战、争议或观点上的分歧”（[1]，第4页），例如“吸烟是否会导致肺癌”（[4]，第107页）。由此可见，他们所理解的“议题”主要指具有公共论辩性质的话题——通常存在明确的正方、反方乃至多方立场，且各方尚未达成共识，从而为讨论和争辩留下空间。尽管议题常被置于政治或伦理的语境下讨论，但实际上，它们广泛渗透于人类思想与活动的各个层面，诸如科学、艺术乃至日常生活。

法乔内也使用“issue”一词，将其作为判断的对象。在定义、技能或倾向方面，他多次强调这一点，例如，他认为具有批判性思维的人应对议题具备清晰认识。（[13]，第3页）在文中后续部分，他经常将“issue”与“controversial”联系使用（[13]，第4页），可见，他所关注的问题类似于论辩议题。除此之外，他并未对“议题”作出明确定义。

杜威就判断的对象提供较为详尽的论述。他区分了“不确定情境”（indeterminate situation）与问题（problem）。问题并非源于主观心理怀疑，而是来自“不确定情境”——这是探究活动的客观起点，体现为个体在与环境互动中遭遇的一种情境性怀疑。（[11]，第105页）这种不确定性使得既有的行为模式或习惯无法有效运作，从而在实践中造成一种“断裂”。值得注意的是，这种不确定性源于情境本身的“弥漫性特质”，而非个人内心的臆测。为了进一步厘清“不确定情境”的边界，杜威它与“怀疑癖”进行了比较：真正的“不确定情境”扎根于具体实践，是客观存在的混乱与失调；而“怀疑癖”缺乏现实情境支撑，是一种无根无据的主观疑惧，属于病态心理状态。（[11]，第106页）

问题的形成是一个动态的过程。不确定情境本身并不直接构成问题，只有当个体意识到情境中的怀疑，并开始尝试将其明确化时，问题才真正形成。例如，在科学实验中，科学家可能最初只是注意到某种异常现象，但通过进一步观察和分析，他们将这种异常转化为一个明确的科学问题，比如“为什么在特定条件下会出现这种现象？”从不确定情景走向问题，实际上是一个认知与选择共同驱动的过程。换言之，问题并非直接呈现在情境中，而是探究者需基于自身的怀疑经验，主动观察情境信息，从而识别出问题。

“问题”与“议题”虽常被视为同类对象，但二者在适用范围与内涵广度上存在显著差异：“议题”通常特指那些具有公共论辩性质、涉及多方立场且尚未达成共识的社会性或规范性话题；相比之下，“问题”的范畴则更为广阔且具有普适性，它涵盖了人类认知与实践的所有领域，既包含需要价值权衡的公共“议题”，也囊括了大量仅涉及事实确认、技术攻关或逻辑推演的非争议性难题。换言之，所有的“议题”在探究初期均可被视为一种特殊的“问题”，但绝非所有的“问题”都能上升为“议题”。一个典型的例证是“如何修复这台损坏的发动机”或“计算某特定函数在 $x=0$ 处的极限值”，它们是标准的“问题”；然而，由于这类任务通常指向唯一的技术解或事实解，不涉及价值观的冲突或多方立场的博弈，缺乏公共论辩的空间，因而它们并不构成通常意义上的“议题”。

对于判断的手段，三个定义文段集合都强调准则⁵（criteria）的作用。准则体现规范性，其是作出判断的根本依据。李普曼区分三类准则：非形式准则（依赖具体情境的信息）、形式准则（特定的制度权威性）与元准则（如相关性、一致性与融贯）。（[21]，第40-41页）学者们的差异主要体现在如何综合利用不同准则以获得最终判断上。⁶ 该问题涉及准则的使用程序或结构，我将其归纳为三种不同类型：并列、论辩及实验。

并列式结构将各项准则横向罗列，却不揭示其层级或相互关系。法乔内将批判性思维拆分为解释、分析、评估、推理、说明和自我调控六项技能即属此类。他明确指出，这些技能的排列并不反映心理、逻辑或认识论上的先后，也不构成教育上的层级结构。（[13]，第11页）他之所以为这种松散结构辩护，是为了避免与其他框架发生冲突、并保留共识空间；但代价是削弱了文段的实践指引力。此外，法乔内在文段中使用“self-regulation”，并强调判断的形成依赖“技能+倾向”。无论是技能还是倾向，都是个体内在的心理属性，这容易造成一种印象：主体仿佛可以独自决定何为合理判断。约翰逊和汉比指出，当下的定义往往忽视了多主体互动中他者的挑战与批判，而真正的判断检验恰是在互动情境中进行的。（[20]，第427页）因此，这里更根本的问题在于：判断的检验尺度究竟是内在的，还是外在的？从文本证据来看，法乔内采取了内在尺度的立场。

拜林和巴特斯比的探究方法体现了论辩式结构。探究过程被理解为一个论辩的、情境敏感的过程，包含六个核心阶段：识别议题（identify the issue）；描绘对立立场；考察语境——包括论辩、知识与历史层面的背景；运用认识论规范——如资料来源的可靠性、因果推理与统计证据；权衡支持与反对理由——运用权衡

⁵ 准则与规范（norm）含义十分接近，但也有细微差别，准则是判断对错、好坏的原则性指导标准，而规范是具体、可执行的行为规则和要求。

⁶ “准则”这个词有时未被直接使用，但在相关的定义文段集合中，我们依然能够找到大量契合李普曼界定的内容。例如，法乔内对技能与倾向的阐述，实质上便是通过确立相应的行动准则展开的；杜威亦然，他针对探究的五个阶段及其涉及的知识，也提出了诸多准则。

推理；得出有理据的判断——并对其确信程度进行评估。（[4]，第 79 页）他们并未描述对论辩议题的具体探究过程或人们的经验状态，但旨在提供一种规范性的指引，以指导人们在论辩场景中，就争议点作出合理判断。为此，他们提出了在探究过程中可能使用的各类准则，包括论证概念及其结构、谬误、论证类型、专家意见以及如何识别议题等。在判断的检验方面，他们强调，个体应主动寻求批判性意见、反对观点以及其他可能的替代论证。（[1]，第 202 页）固然个体可以自行设想反对观点，但真正具有检验力的反对意见往往来自于其他主体在论辩场域中的挑战。由此可见，拜林和巴特斯比并未将检验的尺度置于个体之中，而是采取了外部检验的路径。

杜威则采用一种实验式结构，在该模式下，个体需在与外部环境的互动中不断调整对准则的运用。杜威的探究理论具有时间与功能两个维度：在时间维度上，探究包括三个阶段——最初，不确定情境的出现激发探究的需要，随后，展开主动的探究活动，最终，探究以获得判断为终点，使情境得以确定与统一；在功能维度上，探究过程包含观察、问题的确立、假设的提出、推理与实验五个功能性阶段。（[7]，第 280 页）这些阶段并非线性推进，而是交织、循环、不断修正的动态过程。杜威指出，这一过程充满试错——虚假线索与无效尝试在所难免。（[10]，第 74 页）探究追求正确判断，但也包含从错误中学习的历程。在实验中，我们用客观环境的反馈来检验假设，杜威强调探究起源于情境，而对假设的检验同样需要采取实际行动去改造情境。这里的情境不仅是自然环境，更涉及社会环境，涉及其它人的观点和看法。在人与环境之间的互动，假设可能逐步变成得到保证的断言（warranted assertions），这是成功探究的产物，能作为后续探究的出发点。（[11]，第 8-9 页）判断与断言很接近，但并不相同，判断反映的是内心的结果，而断言则是公开的行为。在探究中，最开始出现的是判断，并且判断的数量不只一个，而是判断的序列，“整个思维过程就在于作出一系列相互关联的判断，这些判断彼此支持，最终导向一个终极判断——即结论”。（[10]，第 118 页）但判断只是内心的想法，它未必得到公开的验证，那些通过外部检验的判断才会成为得到保证的断言，进而解决问题。

5 结论

本文在约翰逊和汉比工作的基础上提出了一个新标准去评价批判性思维定义。该标准以定义文段集合为评价对象，具有三个元问题，并基于文段对元问题的回答区分了内部融贯标准与对比标准。内部融贯标准针对单个定义文段集合，而对比标准则衡量多个不同的文段集合。我选择了三个有代表性的文段集合，并基于新标准中的元问题，对不同文段的内容进行分析。

我得到如下三个结论。第一，三个定义文段集合整体上均符合内部融贯标准：没有发现存在循环论证或前后不一致，在可理解性方面表现突出，能够清晰地向读者呈现其核心观点，并且都对三个元问题作出了显性或隐性的回应。

第二，依据对比标准的规则一，杜威的文段在前两个元问题的回答上更优。他以显性的方式详细阐发了心智的主动操作及对象的被动呈现对于“思维”的影响，由此给出多种与批判性思维相对的思维类型，而法乔内和拜林的定义文段在前两个元问题上均为隐性回答，研究者只能通过其关联词的使用间接推知其观点。

第三，依据对比标准的规则二，杜威的文段在第三个元问题的回答上表现得更为“厚”。在法乔内、拜林和巴特斯比的定义文段中，判断的对象源于论辩议题。就判断的实现手段而言，法乔内的并列式技能结构仅对若干技能进行了罗列，因而其适用范围和指引力最为有限；拜林和巴特斯比的论辩式结构则提供了清晰的程序和丰富的准则，具有较强的指引性。然而，其规范框架不注重心智操作的心理过程，对于如何从不完善过渡到完善欠缺指引。相比之下，杜威的实验式结构以具体情境及怀疑经验为起点，问题由探究者在情境中发现，从而使探究活动与广阔的生活场景紧密结合，包含论辩议题但不局限于论辩议题。更重要的是，杜威将最终检验判断的尺度置于外部环境之中，注重心智操作与环境的交互，并强调在探究过程中对错误的反思、修正及经验的连续性，这为判断从不完善走向完善提供了理论机制，也使其更能适应复杂多变的实践情境。

参考文献

- [1] S. Bailin and M. Battersby, 2015, *Reason in the Balance: An Inquiry Approach to Critical Thinking*, Indianapolis, IN: Hackett Publishing.
- [2] S. Bailin, R. Case, J. R. Coombs and L. B. Daniels, 1999, “Common misconceptions of critical thinking”, *Journal of Curriculum Studies*, **31(3)**: 269–283.
- [3] S. Bailin, R. Case, J. R. Coombs and L. B. Daniels, 1999, “Conceptualizing critical thinking”, *Journal of Curriculum Studies*, **31(3)**: 285–302.
- [4] M. Battersby and S. Bailin, 2018, *Inquiry: A New Paradigm for Critical Thinking*, Vol. 7, Windsor, ON, Canada: University of Windsor.
- [5] J. Bosworth, 2014, “þencan”, *An Anglo-Saxon Dictionary Online*, Edited by T. N. Toller, C. Sean, and O. Tichy. Faculty of Arts, Charles University. <https://bosworthtoller.com/31654>.
- [6] J. Bosworth, 2014, “þyncan”, *An Anglo-Saxon Dictionary Online*, Edited by T. N. Toller, C. Sean, and O. Tichy. Faculty of Arts, Charles University. <https://bosworthtoller.com/32414>.
- [7] M. J. Brown, 2012, “John dewey’s logic of science”, *HOPOS: The Journal of the International Society for the History of Philosophy of Science*, **2(2)**: 258–306.
- [8] J. Dewey, 1896, “The reflex arc concept in psychology”, *Psychological Review*, **3(4)**: 357–370.

- [9] J. Dewey, 1910, *How We Think*, Boston, MA: D. C. Heath & Co.
- [10] J. Dewey, 1933, *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*, Boston, MA: Heath.
- [11] J. Dewey, 1938, *Logic: The Theory of Inquiry*, New York, NY: Henry Holt and Company.
- [12] P. A. Facione, 1990, "Critical thinking: a statement of expert consensus for purposes of educational assessment and instruction (the delphi report)", Millbrae, CA: California State University. <https://eric.ed.gov/?id=ED315423>.
- [13] P. A. Facione, 1990, *Executive Summary of "The Delphi Report"*, Millbrae, CA: The California Academic Press.
- [14] C. Geertz, 1973, "Thick description: Toward an interpretive theory of culture", in C. Geertz(ed.), *The Interpretation of Cultures: Selected Essays*, pp. 311–323, New York, NY: Basic Books.
- [15] H. P. Grice, 1975, "Logic and conversation", in P. Cole and J. L. Morgan(eds.), *Syntax and Semantics: Vol. 3. Speech Acts*, pp. 41–58, New York, NY: Academic Press.
- [16] D. Harper, 2025, "Etymology of think", *Online Etymology Dictionary*, Retrieved October 30, 2025, <https://www.etymonline.com/word/think>.
- [17] D. Hitchcock, 2022, "Critical thinking", in E. N. Zalta and U. Nodelman(eds.), *The Stanford Encyclopedia of Philosophy*, Winter 2022 Edition, <https://plato.stanford.edu/archives/win2022/entries/critical-thinking/>.
- [18] T. Hoy, 2023, "John dewey", in E. N. Zalta(ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2023 Edition, <https://plato.stanford.edu/archives/fall2023/entries/dewey/>.
- [19] R. Johnson, 1992, "The problem of defining critical thinking", in S. P. Norris(ed.), *The Generalizability of Critical Thinking*, pp. 39–53, New York, NY: Teachers College Press.
- [20] R. Johnson and B. Hamby, 2015, "A meta-level approach to the problem of defining 'critical thinking'", *Argumentation*, **29(4)**: 417–430.
- [21] M. Lipman, 1987, "Critical thinking: What can it be?", *Analytic Teaching*, **8(1)**: 38–43.
- [22] Oxford University Press, "Think, v.1, etymology", *Oxford English Dictionary*, Retrieved November 1, 2025, <https://doi.org/10.1093/OED/2211168647>.
- [23] F. Riga, 2020, "Pragmatism—John Dewey", in B. Akpan and T. Kennedy(eds.), *Science Education in Theory and Practice*, pp. 227–239, Cham, Switzerland: Springer.
- [24] G. Ryle, 1968, "The thinking of thoughts: What is 'Le Penseur' doing?", *Collected Papers, Volume 2: Collected Essays 1929–1968*, pp. 494–510, Abingdon: Routledge.
- [25] I. Scheffler, 1960, *The Language of Education*, Springfield: Charles C. Thomas.
- [26] B. Williams, 1985, *Ethics and the Limits of Philosophy*, Cambridge, MA: Harvard University Press.

A Criteria to Evaluate Critical Thinking's Definitions and Its Application

Runcheng Liang

Abstract

Scholars have proposed numerous definitions of critical thinking, and choosing which definition to adopt is a contested issue. Johnson and Hamby also recognized this, taking an alternative approach by examining definitions of critical thinking from a meta-level perspective and attempting to propose a set of standards for evaluating definitions. This paper similarly adopts a meta-theoretical perspective but seeks to refine the standards proposed by Johnson and Hamby and apply them to the evaluation of critical thinking definitions. Specifically, their standards focus on individual definitional items, neglecting the explanatory passages provided by the definers, giving insufficient attention to the discussion of “thinking” within the collection of definitional passages, and failing to offer criteria for comparing different collections of definitional passages. To address this, this paper derives three meta-questions from the etymology of the terms “critical” and “thinking” in English, and, based on the theoretical possibilities of the definitional passages’ responses to these meta-questions, proposes an internal coherence standard and a comparative standard. Finally, the new standards are applied to the collections of definitional passages by Dewey, Facione, and Bailin. The results show that all three meet the internal coherence standard, while in comparison, Dewey’s passages provide explicit answers regarding “thinking” and the comparison of thinking types, and offer greater “thickness” in the exposition of “critical”, therefore, it has greater practical guiding power.