

CONNECTICUT INSURANCE LAW JOURNAL

VOLUME 25

2018-2019

NUMBER 2

CONTENTS

ARTICLES

UNDERSTANDING THE “EXHAUSTION
OF COVERAGE” DOCTRINE IN THE
CONTEXT OF CONTINUOUS TRIGGER
COVERAGE

James M. Fischer

263

TRANSITIONING TO A NEW VIEW:
COMING TO SEE HEALTH INSURANCE
COVERAGE FOR GENDER DYSPHORIA
IN A NEW LIGHT

Robert A. Whitney

297

PLAIN MEANING, EXTRINSIC
EVIDENCE, AND AMBIGUITY: MYTH
AND REALITY IN INSURANCE POLICY
INTERPRETATION

Kenneth S. Abraham

329

UNDERSTANDING THE “EXHAUSTION OF COVERAGE” DOCTRINE IN THE CONTEXT OF CONTINUOUS TRIGGER COVERAGE

JAMES M. FISCHER*

Excess insurance provides additional coverage beyond that provided by the underlying first layer of insurance, usually referred to as primary insurance policies. Excess insurance may, in turn, be layered with multiple layers of excess coverage. The combined, aggregate insurance policies are usually referred to as the “insured’s coverage program.” A fundamental issue with respect to the relationship between excess and underlying insurance is determining the event that triggers the specific excess coverage and brings it into play. Traditionally, the triggering event has been the “exhaustion” of the underlying insurance policy(ies), whether the underlying policy is primary or excess.

The Exhaustion requirement was initially developed when insurance coverage was siloed within distinct policy periods.¹ Siloing means there is a vertical tower of coverage, with layers of coverage, beginning with primary coverage² and extending upwards through one or multiple layers of excess coverage. Exhaustion envisions that as each underlying layer of coverage is used up, (i.e., the policy limits are expended (“exhausted”)), the next layer of coverage can be accessed. While this approach is not without problems, it has worked tolerably well as a means of regulating the relationship between insurers who provided underlying and overlaying levels of coverage within a single policy period.

The advent of continuous trigger coverage requires a reassessment of the relationship between underlying and overlaying insurers because siloing of coverage within a single policy period is no longer the norm.

* Professor of Law, Southwestern Law School, Los Angeles, California.

¹ Initially, courts simply referred to the exhaustion of “all primary coverage” before an insured could access overlaying excess coverage. *See, e.g., Olympic Ins. Co. v. Emp’rs Surplus Lines Ins. Co.*, 178 Cal. Rptr. 908, 912 (Cal. Ct. App. 1981). The term “horizontal exhaustion” seems to have first appeared in an Illinois Court of Appeal decision in 1994. *See United State Gypsum Co. v. Admiral Ins. Co.*, 643 N.E.2d 1226, 1261 (Ill. App. Ct. 1994) (distinguishing between “horizontal” and “vertical” exhaustion of underlying primary insurance coverage).

² Primary coverage refers to the first layer, often called the “working layer,” of coverage. Most claims against an insured are resolved within this first layer of coverage.

Rather, insurance coverage extends to losses than run across multiple policy periods.³ Insured losses can no longer be compartmentalized within a single policy period; rather, insurers (primary and excess) find that coverage is triggered whenever any portion of a loss can be deemed to have occurred in a policy period.⁴ The aggregation of coverage accomplished by the use of a continuous trigger led the California Supreme Court to characterize the resulting aggregation of coverage across multiple policy periods as “one

³ A continuous trigger means that an ongoing event, such as a pollutant contaminating underground water reservoirs, may trigger multiple policies that provide coverage during the contamination period. Courts applying a continuous trigger treat the injury attributable to the insured event as indivisible. *State v. Cont’l Ins. Co.*, 281 P.3d 1000 (Cal. 2012). If the injury can be confined to a particular policy period, courts do not apply a continuous trigger. *See, e.g., In re Silicone Implant Ins. Coverage Litig.*, 667 N.W.2d 405, 419 (Minn. 2003). *Cf. N. States Power Co. v. Fid. & Cas. Co. of N.Y.*, 523 N.W.2d 657, 662-63 (Minn. 1994) (adopting continuous trigger because of the “scientific complexity of the issues involved, the extended period of time over which damages may have occurred before discovery, and the number of parties potentially involved”), *and* RESTATEMENT OF THE LAW OF LIAB. INS. §44, cmt. b (AM. LAW INST., Tentative Draft No. 1, 2016) (“For liability claims involving divisible harm, courts generally will attempt to allocate among the policy periods according to the actual injury or harm that occurred during the policy period even if the total harm occurred over a long period of time.”).

⁴ Most jurisdictions, including California, use the “injury in fact” test to implement the continuous trigger. *See Montrose Chem. Corp. v. Admiral Ins. Co.*, 913 P.2d 878, 894 (Cal. 1995) (“Under an injury-in-fact trigger, coverage is first triggered at that point in time at which an actual injury can be shown, retrospectively, to have been first suffered. This rationale places the injury- in-fact somewhere between the exposure, which is considered the initiating cause of the disease or bodily injury, and the manifestation of symptoms, which, logically, is only possible when an injury already exists. In the context of continuous or progressively deteriorating injuries, the injury-in-fact trigger, like the continuous injury, affords coverage for continuing or progressive injuries occurring during successive policy periods subsequent to the established date of the initial injury-in-fact.” (citations omitted)).

giant ‘uber-policy’ with a coverage limit equal to the sum of all purchased insurance policies across the period of loss.”⁵

Treating policies within a continuous trigger period as an “Uber Policy” does not, however, directly address how overlaying policies should respond to a loss relative to the responsibility of an underlying policy, particularly that of primary insurers. In adopting the “continuous trigger” and “all sums stacking” approaches, the California Supreme Court did not address how coverages within the “uber-policy” should be sequenced. This is, however, a fundamentally important question because the “uber-policy” is an illusion if it cannot be effectively and efficiently accessed by the insured. As will be addressed in this paper, the sequencing of coverage provided by horizontal exhaustion results in the irony that the more coverage the insured purchases, the greater the likelihood much of the purchased insurance will not be available to pay claims put in coverage under the “continuous trigger” “all sums stacking” doctrines.

For the most part, insurers, as a group have argued for “horizontal” exhaustion to determine when overlaying coverages first become responsible for a loss. Horizontal exhaustion requires that all underlying layers of coverage be “exhausted” before an overlaying level of coverage must provide coverage.⁶ For example, look at Schematic 1, on page 266. Assume an insured event triggers the Alpha, Beta, and Gamma policies under a continuous trigger. Under the horizontal exhaustion test, Excess Insurer Omega would argue that all underlying insurance from primary insurers Alpha, Beta, and Gamma would have to be exhausted before Omega would have a coverage obligation. Thus, even if the underlying Gamma policy was exhausted, Omega would have no coverage obligation until the underlying Alpha and Beta policies were also exhausted. Under this approach, second layer excess Insurer Sigma would argue that it has no coverage obligations

⁵ *State v. Continental Ins. Co.*, 281 P.3d 1000, 1008 (Cal. 2012). In this decision, the California Supreme Court used the term to describe the practical effect on insurers who provide coverage over a period of time to an insured who has sustained a progressive loss (e.g., environmental contamination) and coverage obligations are subject to both a “continuous trigger” and “all sums stacking.” The practical effect of adopting a “continuous trigger” and “all sums stacking” is that a loss that extends over several years separately triggers each year’s policies cumulatively. *Id.*

⁶ *See., e.g., Emp’rs Ins. Co. of Wausau v. Granite State Ins. Co.*, 330 F.3d 1214, 1221 (9th Cir. 2003).

until all the underlying coverage provided by Primary Insurers Alpha, Beta, Gamma, and First Layer insurers Epsilon, Theta, and Omega was exhausted. Insurers often use the metaphor of a “rising tide” to illustrate the manner in which layers of coverage within an insured’s coverage program are accessed under the horizontal exhaustion approach.

Schematic 1

2d Layer, Excess Insurers	Psi	Sigma	Chi
1st Layer, Excess Insurers	Epsilon	Theta	Omega
Primary Layer, Insurers	Alpha	Beta	Gamma
	Year 1 Policies	Year 2 Policies	Year 3 Policies

Policyholders, on the other hand, usually advance the argument of “vertical” exhaustion.⁷ Vertical exhaustion applies traditional siloing to continuous trigger coverage cases. Under vertical exhaustion, overlaying insurer coverage is only dependent on the exhaustion of the specific policy underlying the overlaying insurance policy. For example, under Schematic 1, using vertical exhaustion, once the underlying Gamma policy was exhausted, Omega’s excess policy would be triggered, even though neither the Alpha nor the Beta policies were exhausted.⁸ Vertical exhaustion is consistent with the traditional approach used when there is no continuous trigger. Each policy year is siloed (kept apart) from other policy years. Absent a continuous trigger, an insured event in year 2 would only trigger the Beta primary policy. The Alpha and Gamma policies would not be

⁷ *Montrose Chem. Corp. v. Superior Court*, 222 Cal. Rptr. 3d 748, 756 (Ct. App. 2017), *rev. granted* *Montrose Chem. Corp. v. S.C. (Canadian Universal Ins. Co.)* 406 P.3d 327 (Cal. 2017) (mem.).

⁸ *See, e.g., Outboard Marine Corp. v. Liberty Mut. Ins. Co.*, 670 N.E.2d 740, 748 (Ill. App. Ct. 1996). This approach is particularly helpful to insureds if coverage is uneven because horizontal exhaustion will preclude immediate access to coverage in policy years with more excess coverage until underlying coverage in other policy years (with possibly lesser coverage (in terms of policy limits)) is exhausted.

triggered. If the Beta policy was exhausted, only the Theta policy would respond, and the Sigma policy would only respond if the Theta policy was exhausted. None of the excess insurers in years 1 and 3 would have coverage obligations upon the exhaustion of the Beta policy, or any year 2 policies for that matter. If a continuous trigger is applied, under vertical exhaustion, once the Beta policy is exhausted, the Theta policy is triggered. This triggering occurs even though the Alpha and Gamma policies are not exhausted.

In this paper I explore a number of issues that relate to the selection of horizontal or vertical exhaustion in the context of a continuous trigger approach to coverage. The presence of one or more of these issues will ordinarily result in the inability to completely exhaust a successive layer of coverage, for example, the first layer of excess coverage shown on Schematic 1 across policy year 1 through 3. When this occurs, under the theory of horizontal exhaustion all overlaying policies anywhere in the coverage program will now escape any obligation to provide coverage. For example, if in Schematic 1 the Beta policy was not exhausted, under horizontal exhaustion all the overlaying excess insurers in the first and second excess layers of coverage for policy years 1-3 would escape coverage obligations. The consequence will be, if horizontal exhaustion is adopted, a cascading loss of coverage throughout all the towers of coverage from that point onward.

The first part of this paper addresses whether policy language generally calls for adoption of horizontal or vertical exhaustion. Should the selection of horizontal or vertical exhaustion be a rule or interpretation or a rule of construction, like the doctrine of construction against the drafter (*contra proferentem*)? If selection of the exhaustion rule is seen as one of the policy interpretations, what language should be understood as selecting one exhaustion approach over the other?

This paper next examines a number of doctrines that apply to the exhaustion issue in general.

These include:

First, should a settlement with an underlying insurer affect the obligations of higher layer insurers when they fail to exhaust policy limits?

Second, and related to Second, should non-accumulation and prior insurance provisions affect the exhaustion issue?

Third, should the insolvency of an underlying insurer affect the exhaustion issue?

Fourth, should loss allocation agreements between policyholders, such as indemnity and hold harmless agreements, affect the exhaustion issue?

This paper concludes with an argument that when courts impose coverage obligations under a “continuous trigger” “all sums stacking” approach the obligations of underlying and overlaying insurers, whose policies are triggered by the “continuous trigger” theory, should be determined under a vertical rather than horizontal exhaustion approach. Adoption of horizontal exhaustion is not only inconsistent with the concept of the “giant uber-policy”, adoption of horizontal exhaustion puts of the insured’s entire coverage program at risk whenever any part of the program fails.

I. JUSTIFICATION FOR HORIZONTAL OR VERTICAL EXHAUSTION BASED ON POLICY LANGUAGE

Courts justify the application of either horizontal exhaustion or vertical exhaustion on two grounds: policy language and public policy; however, in this context the differences between the two grounds are not always well defined. Moreover, courts have not always been consistent in the treatment of policy language as supporting horizontal as opposed to vertical exhaustion.

Some, but not all, courts have adopted vertical exhaustion when an overlaying level of coverage specifically identifies an underlying policy which has been exhausted. For example, in *Viking Pump* the New York Court of Appeals held that vertical exhaustion was called for when the overlaying policy referenced underlying policies as those “listed as an underlying policy in the declarations.”⁹ In that context, the overlaying policy was triggered upon the exhaustion of the described referenced policies without regard to the status of the other policies at the same layer of any of the referenced policies.

Similarly, in *State v. Continental Insurance Co.*, the appellate court stated that language in an excess insurance policy that liability attached upon an “Ultimate Net Loss” which referenced the specified retention “seems to be the very definition of vertical exhaustion.”¹⁰

On the other hand, language that states the overlaying coverage is excess of [referenced policies] “and any other underlying insurance

⁹ *In re Viking Pump*, 52 N.E.3d 1144, 1157 (N.Y. 2016).

¹⁰ 223 Cal. Rptr. 3d 716, 727 (Ct. App. 2017).

providing coverage to the insured” has been interpreted as calling for horizontal exhaustion.¹¹

In *Montrose Chemical Corp.* the court held that overlaying coverage that referenced specific coverage required horizontal exhaustion when the policy language specifically referenced a provision that incorporated “other underlying insurance.”¹² The policy language, as restated by the court, provided:

[T]he insurer agrees to pay on behalf of the insured the ultimate net loss in excess of the retained limit hereinafter stated.” The declarations then identify the underlying policies to which the American Centennial policies are specifically in excess (the “scheduled policies”).¹³

The court noted the “Retained Limit” clause referred to the overlaying insurer’s liability as “excess of the identified underlying insurance and the applicable limits of any other underlying insurance collectible by the insured.”¹⁴ While the court could have treated the conflicting language as ambiguous and resolved the conflict in the insured’s favor,¹⁵ the court chose to emphasize the broader language in the Retained

¹¹ *Montrose Chem. Corp.*, 222 Cal. Rptr. 3d at 763 (discussing the American Centennial excess policies).

¹² *Id.*

¹³ *Id.*

¹⁴ *Id.* at 764 (“The ‘retained limit’ clause: This clause provides: ‘[T]he company’s liability shall be only for the Ultimate net loss in excess of the insured’s retained limit defined as the greater of: [...]the total of the applicable limits of the underlying policies listed in the [declarations] hereof, and applicable limits of any other underlying insurance collectible by the insured.’ (italics added.) This clause thus expressly states that the excess insurer’s liability is in excess of the identified insurance and the applicable limits of any other underlying insurance collectible by the insured.”).

¹⁵ This is the rule of *Contra Proferentum*, or construction against the drafter. Often, but not always, courts will find that an inconsistency between insurance provisions in the same policy create ambiguity. That ambiguity is then resolved in the favor of the policyholder if a reasonable interpretation of the ambiguity so permits.

Limits clause as supporting horizontal exhaustion because of the general reference to “other underlying insurance.”¹⁶

The *Montrose Chemical Corp.* also found support for adoption of horizontal exhaustion when the overlaying policy references specific underlying coverage by looking at the overlaying policy’s “Other Insurance” clause, which provided that the overlaying policy was excess to “both scheduled and unscheduled policies.”¹⁷ However, this use of the “Other Insurance” clause language to support horizontal exclusion has been rejected by other courts as improper for several reasons. First, “Other Insurance” clauses are seen by these courts as a means of allocating responsibility for a loss among insurers, not as a means of avoiding or deflecting liability to an insured.¹⁸ When the “continuous trigger” and “all sums stacking” approach is adopted, the insured’s entire coverage program is available to provide compensation for the loss. Allowing insurers to use the “Other Insurance” provision to escape or deflect their coverage obligations for a loss sustained by their common insured is inconsistent with the view that “Other Insurance” provisions are intended to prevent payments to the insured in excess of the

¹⁶ *Montrose Chem. Corp.*, 222 Cal. Rptr. 3d at 764 (italics in the original).

¹⁷ *Id.* (“The ‘other insurance’ clause : This clause states: “ ‘If other collectible insurance...is available to the insured covering a loss also covered hereunder (except insurance purchased to apply in excess of the sum of the retained limit and the limit of liability hereunder) the insurance hereunder shall be in excess of and not contribute with, such other insurance.’ ” This clause thus provides that the American Centennial policies are excess to both *scheduled* and *unscheduled* policies.).

¹⁸ See, e.g., *Continental Ins. Co.*, *supra* note 5, 223 Cal. Rptr. 3d at 727 (collecting California decisions); *cf.* *Shade Foods, Inc. v. Innovative Product Sales & Marketing, Inc.*, 93 Cal. Rptr.2d 364, 407 (Ct. App. 2000):

The other-insurance clause, as we have seen, does not excuse the insurer from discharging its independent obligation to indemnify the insured up to policy limits, though it gives the insurer a right to an adjudication allocating the indemnity obligation between it and the other insurer.

The court added that the insurer’s use of the “Other Insurance” clause to deflect its obligations will support a jury’s determination of ‘bad faith.’”

loss, consistent with the principle of indemnity.¹⁹ When the “Other Insurance” clause purports to require exhaustion of other available insurance before the policy may be accessed, the “Other Insurance” clause is typically referred to as an “excess” or “escape” type provision. “Escape” and “Excess” type provisions create difficulties because the other available insurance policy(ies) may also have “escape” or “excess” type “Other Insurance” clauses, which results in mutual repugnancy.²⁰

Second, and more generally accepted, many courts limit the use of “Other Insurance” clauses to insurers providing the same layer or level of coverage.²¹ Using the “Other Insurance” clause as a justification for horizontal exclusion is not consistent with this limitation because horizontal exhaustion requires exhaustion of underlying layers of coverage, which are clearly not at the same layer or level of coverage as the policy containing the “Other Insurance” clause.²² Courts relying on the “other Insurance” clause

¹⁹ See ROBERT E. KEETON, ALAN I. WIDISS & JAMES M. FISCHER, INSURANCE LAW § 3.11, at 217 (2d ed. 2016) (“Duplication of coverage raises the prospect that the indemnity principle will be violated. Other Insurance Provisions seek to ameliorate duplicate coverage so that indemnity principle is preserved.”) (footnote omitted).

²⁰ *Dart Indus., Inc. v. Commercial Union Ins. Co.*, 52 P.3d 79, 93 (Cal. 2002) (“[P]ublic policy disfavors ‘escape’ clauses, whereby coverage purports to evaporate in the presence of other insurance. This disfavor should also apply, to a lesser extent, to excess-only clauses, by which carriers seek exculpation whenever the loss falls within another carrier’s policy limit.”)

²¹ *In re Viking Pump*, 52 N.E.3d 1144, 1157 (N.Y. 2016) (“[W]e stated in *Consolidated Edison* that ‘other insurance’ clauses’ apply when two or more policies provide coverage during the same period, and they serve to prevent multiple recoveries from such policies, “and that such clauses ‘have nothing to do’ with whether any coverage potentially exist[s] at all among certain high-level policies that were in force during successive years.”)(citation omitted).

²² See KEETON, WIDISS & FISCHER, *supra* note 19, at 220. See also *Dart Indus., Inc. v. Commercial Union Ins. Co.*, 52 P.3d 79, 93 (Cal. 2002) (“[H]istorically, ‘other insurance’ clauses were designed to prevent multiple recoveries when more than one policy provided coverage for a particular loss. On the other hand, ‘other insurance’ clauses that attempt to shift the burden away from one primary insurer wholly or largely to other insurers have been the objects of judicial distrust. Public policy disfavors

as supporting the adoption of horizontal exhaustion have not explained affirmatively why the “Other Insurance” clause should be applied to different layers of insurance.²³

Many courts have stated that a particular form of exhaustion is required based on precedent.²⁴ This argument, however, begs the question as to why the initial decision that constitutes the precedent was made in the first place. Horizontal exhaustion has been justified as consistent with the price differential between primary and excess coverage.²⁵ This appears, however, to more an assertion than a justification. Price reflects the actuarial assessment that the covered risk will occur. Unless we know the underlying actuarial assessments, we cannot determine what the price specifically envisions as to risk.²⁶ Some courts have justified adoption of horizontal exhaustion as necessary to prevent insureds from manipulating the sources of recovery and ignoring the distinctions between primary and excess insurance.²⁷ Again, these appear to be more in the nature of assertions rather than reasoned arguments as courts making these arguments have not identified instances of manipulation nor reasons why horizontal exhaustion

escape clauses, whereby coverage purports to evaporate in the presence of other insurance. This disfavor should also apply to a lesser extent, to excess-only clauses, by which carriers seek exculpation whenever the loss falls within another carrier’s policy limit. Partly for this reason, the modern trend is to require equitable contributions on a pro rate basis from all primary insurers regardless of the type of ‘other insurance’ clause in their policies.”(citations omitted)).

²³ *E.g., Montrose Chem. Corp. v. Superior Court*, 222 Cal. Rptr. 3d 748, 767 (Ct. App. 2017) (distinguishing a prior decision that contained language that “Other Insurance” clauses only applied to allocations among insurers and should not be used to deflect insurer coverage obligations). Distinguishing a prior decision is not, however, the same as affirmatively demonstrating the soundness of the position reached.

²⁴ *See Iolab Corp. v. Seaboard Sur. Co.*, 15 F.3d 1500, 1504 (9th Cir. 1994).

²⁵ *See Mo. Pac. R.R. Co. v. Int’l Ins. Co.*, 679 N.E.2d 801, 809 (Ill. Ct. App. 1997).

²⁶ Applying horizontal exhaustion to primary layers of coverage may be justified because of the defense obligation that attaches at the primary level and which is reflected in the risk assumed by primary insurers. *See infra* notes 43-48 and accompanying text.

²⁷ *See, e.g., U.S. Gypsum Co. v. Admiral Ins. Co.*, 643 N.E.2d 1226, 1261 (Ill. Ct. App. 1994).

is more congruent with the primary-excess distinction than vertical exhaustion.

The claim has been made that vertical exhaustion should be adopted because it is most consistent with the “All Sums” method of allocation that has been adopted by most jurisdictions to address coverage obligations under the continuous trigger doctrine.²⁸ While courts have not explained why vertical exhaustion is most consistent with the “All Sums” approach,²⁹ an explanation may lie in the third reason courts have given for adopting a particular exhaustion approach for reasons other than policy language. That reason is complexity. It becomes exceedingly difficult to determine how overlaying coverage should be accessed when the insurance plan extends over a lengthy period of time, and the doctrine in the jurisdiction holds all that policies have been triggered – creating the so-called “Uber” policy.³⁰ Consider for example the problem of uneven layers of insurance described in Schematic 2, on page 274, where some policy periods have 7 layers of coverage; other policy periods have 5 layers of coverage, etc. More importantly, the layers of coverage have different limits. For example, in year 3, Insurer J has \$19 million of coverage over an underlying primary layer, provided by Insurer L, of \$1 million. In year 5, two successive insurers,

²⁸ See *Viking Pump*, 52 N.E.3d at 1156 (collecting decisions); see also RESTATEMENT OF THE LAW OF LIAB. INS., §§ 42 cmt. c, 44 cmt. c (Proposed Final Draft 2017).

²⁹ In some respects, the problem may be one of framing. For example, in Schematic 1, should the insurance provided by the nine insurers be seen as coverage (singular) or as coverages (plural)? If the entire program of insurance is seen as an integrated package of insurance (the “uber” policy), it may be easier to visualize coverage as rising evenly from bottom to top (horizontal exhaustion) than spiking as individual coverages are exhausted (vertical exhaustion).

³⁰ Courts that adopt horizontal exhaustion tend to put this concern in the “never mind” category. See *Montrose Chem. Corp.*, 22 Cal. Rptr.3d at 1335-36 (“Montrose argues finally that mandatory horizontal exhaustion is ‘unworkable in practice’ because of the complexity of its coverage portfolio. We do not doubt that allocating more than \$200 million in liability across more than 100 policies covering nearly 25 years is likely to be a complicated process. The complexity, however, is not relevant to our analysis, as we cannot, in the service of expediency, impose obligations that are inconsistent with the terms of the contracts Montrose itself negotiated.”).

Insurers S and T, provided the same amount of coverage over an underlying layer of coverage, provided by Insurer R. How would Insurer S and T's obligations be triggered? Is Insurer T a second layer or third layer insurer for purposes of horizontal exhaustion? If overlaying policies are simply triggered as underlying dollar amounts, e.g., \$1 million, \$5 million, \$10 million, etc. are reached, the internal structure of coverage within each policy year is compromised. If each layer within a policy period must await the exhaustion of the underlying layer, regardless of total losses within the triggered policy periods, the metaphor of a rising tide is an illusion.³¹

Schematic 2

\$50 Million				\$50 Million			
\$40 Million				Insurer P		Insurer CC	\$40 Million
\$30 Million					Insurer U	Insurer BB	\$30 Million
\$20 Million		Insurer H	Insurer K			Insurer AA	\$20 Million
\$15 Million	Insurer D	Insurer G	Insurer J	Insurer O	Insurer T	Insurer Z	\$15 Million
\$10 Million						Insurer Y	\$10 Million
\$5 Million	Insurer C			Insurer N	Insurer S	Insurer X	\$5 Million
\$1 Million	Insurer B	Insurer F	Insurer J			Insurer W	\$1 Million
	Insurer A	Insurer E	Insurer L	Insurer M	Insurer R	Insurer V	
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	

Another problem with implementing horizontal exhaustion is that policy limits within the various layers of the insured's coverage program may be uneven from year to year. For example, in year 1, the 4th layer of

³¹ Another example of this concern is provided by *Westport Ins. Corp. v. Appleton Papers, Inc.* 787 N.W. 2d 894, 918-19 (Wis. Ct. App. 2010).

coverage provided by Insurer D attaches at \$10 million; in year 4 the 4th layer of coverage attaches at \$40 million. Similarly, in years 1, 2, 3, 5, and 6, the primary layer provides \$1 million in coverage, but in year 4, the coverage provided by Insurer M is \$5 million. Rather than coverage being distributed evenly by layer and amount through the 6-year coverage period, insurance coverage is unevenly distributed both as to the number of layers of coverage, the amounts of insurance coverage provided in the individual policies, and the aggregate amount of insurance provided in each policy year. Exhausting underlying coverage horizontally across uneven towers of coverage can prove to be a daunting task.³²

A second reason for rejecting horizontal exhaustion on policy ground is that it contains a disguised forfeiture feature: The more coverage an insured purchases, the greater the risk much of the coverage will be lost due to the inability to fully exhaust one of the policies in the coverage program.

Consider for example, the coverage program set out in Schematic 2. If the insured only purchases coverage for year 1, the year 1 coverage is defined by the coverages provided by Insurers A, B, C, and D. If, however, the insured acquires coverage in year 2, the insured’s ability to access the coverage provided by Insurers B, C, and D is now controlled by the need to exhaust the coverages provided by Insurers E, F, G, and H. The problem continues as the insured continues the coverage program into year 6. Now the coverage provided by Insurers B, C, and D is subject to being lost if the insured does not fully exhaust the coverage provided by Insurers V, W, X, Y, Z, etc. Under the horizontal exhaustion approach the failure to exhaust the policy coverage provided by Insurer V puts all of the coverage for all of the years at risk. That is a forfeiture, plain and simple. There is no sound reason why coverage in place (e.g., year 1) should be subject to subsequent events (e.g., coverage placed in year 6). Courts that have adopted horizontal exhaustion have assumed all they are resolving is how the sequencing of policies in place will be accessed. These courts have ignored, or failed to fully consider, the significant likelihood that a policy will not be exhausted and how that failure will affect the insured’s ability to access other triggered policies. And as shown in Parts II through V of this paper, the likelihood that

³² Insurance towers of coverage may be extensive. Appendix A to this illustrates the insurance coverage program of an insured that was involved in coverage litigation involving a toxic waste site. *See State v. Cont’l Ins. Co.*, 15 Cal. App. 5th 1017, 1047-48 (Cal. Ct. App. 2017).

a policy in the insured's coverage program will not be fully exhausted is a real and present danger. In effect, the horizontal exhaustion doctrine operates as a hidden bomb that when detonated by the inability to exhaust a single policy, compromised the entire coverage program.³³ Horizontal exhaustion kills coverage while preserving the illusion of coverage. The unappreciated consequence of horizontal exhaustion is that the more coverage an insured obtains, the less likely it becomes that the insurance purchased will be available to the insured when a loss occurs.

II. SETTLEMENT AND EXHAUSTION OF UNDERLYING LIMITS

A recurring issue today is the effect of a settlement between an insured and an underlying insurer on the overlaying insurer when the settlement is for less than the underlying insurer's policy limits. For example, in Schematic 2, assume Insurer M, with \$5 million in limits, disputes coverage of the claim and the Insured and Insurer M agree to settle the dispute with a payment by Insurer M of \$4 million. What affect, if any, does the settlement have on the coverage obligations of Insurers A through CC?

When an insured settles a coverage dispute with an underlying insurer for less than the policy limits, overlaying insurers will claim that a condition of their coverage obligations – exhaustion by payment of policy limits of underlying coverage – has not been met. Excess insurers will also claim that payment or credit by the insured of any difference between the settlement amount and the policy limits will not satisfy the coverage condition in the overlaying policy, absent express policy language permitting the insured to cover.

Not surprisingly, courts have differed whether the excess insurers' position should be accepted. Many decisions today adopt a formal "follow the policy language" approach. Under this approach, the policy language determines whether the insured may access the excess insurance coverage by covering the gap between the settlement amount and the policy

³³ Forfeiture of policy benefits is disfavored, and courts will generally construe policy terms and adopt rules of policy construction that avoid forfeiture of policy benefits. *See* COUCH ON INSURANCE §§ 22:34-35 (3d ed. 2017). *See also* *Richmond v. Dart Indus, Inc.*, Cal 629 P.2d 23, where the California Supreme Court comments that coverage escapism by insurer use of "Other Insurance" clauses is disfavored. *See supra* note 20. The same sentiments apply to coverage escapism produced by horizontal exhaustion.

limits.³⁴ As, however, with almost all insurance claims, disputes may arise whether the policy language does permit the insured to access an overlaying level of coverage by paying the differences between the policy limits and the settlement amount.

Consider, for example, policy language that simply requires exhaustion of an underlying policy by “actual payment.” Must the payment be made by the insurer in satisfaction of judgment or settlement, or may it be made by the insured? Again, courts have disagreed on this point.³⁵

When relying exclusively on policy language, very slight differences in policy language may result in a loss or preservation of excess insurance coverage when an insured settles a coverage dispute with an underlying insurer for less than the policy limits. For example, in *Zeig v. Massachusetts Bonding & Ins. Co.*,³⁶ the court held that language in an excess insurance policy that conditioned access when the underlying policy was “exhausted in the payment of claims to the full amount of the expressed limits” permitted the insured to satisfy the requirement by a bridge payment.³⁷ In *Ali v. Federal*

³⁴ In some situation, the policy language explicitly permits the insured to cover the gap. See, e.g., Axis Excess D & O Policy, XS 0001 12 10:

This policy shall provide insurance excess of the Underlying Insurance. Liability shall attach to the Insurer only after (i) the Insurers of the Underlying Insurance, the Insureds or others on behalf of the Insureds shall have paid in legal currency amounts covered under the respective Underlying Insurance equal to the full amount of the Underlying Limit...

³⁵ Cf. *Forest Labs. Inc. v. Arch Ins. Co.*, 953 N.Y.S. 2d 460 (N.Y. Sup. Ct. 2012) (holding that “actual payment” language unambiguously requires payment by insurer to exhaust limits); *with Maximus, Inc. v. Twin City Fire Ins. Co.*, 856 F. Supp. 2d 797 (E.D. Va. 2012) (holding that “actual payment” language was ambiguous and could be reasonably construed to permit bridging payment by insured to satisfy exhaustion requirement). By “bridging payment” I mean that the insured assumes the obligation to pay the difference between the policy limits and the amount of the settlement.

³⁶ *Zeig v. Mass. Bonding & Ins. Co.*, 23 F. 2d 665 (2d Cir. 1928).

³⁷ *Id.* at 666. *Zieg* involved a 1st party property insurance policy, which did not contain a duty to defend. A number of courts have distinguished *Zeig* on this ground and have refused to apply *Zieg* to 3d party liability insurance

Ins. Co., however, the court held that language in the excess insurance policy that conditioned access to the policy “only after...all Underlying Insurance has been exhausted by payment of claims...solely as a result of payment of losses thereunder” did not permit the insured to use a bridge payment to satisfy the condition of access to the excess insurance policy.³⁸

In many cases today, the policy language is quite clear that exhaustion must be accomplished by payments by the insurer.³⁹ Of course, a consequence of treating payment by the insurer as the exclusive method of exhausting underlying coverage is that the insured loses overlaying insurance by settling a coverage dispute with the underlying insurer. A number of courts and commentators have argued that such a result imposes a forfeiture on the insured that is violative of public policy. Under this view, when the insured in good faith settles a coverage dispute with the underlying insurer for less than policy limits, that settlement should not cause the insured to lose the excess insurance in place when the overlaying insurers are not prejudiced.⁴⁰ Alternatively, if the dispute is not resolved and the insurer prevails on the coverage dispute, that specific policy in the tower of coverage is not exhausted and, under the horizontal exhaustion approach, the entire coverage program collapses from that point upward.

The problem is compounded in the continuous trigger context when the issue is whether the court should apply either horizontal exhaustion or vertical exhaustion. If vertical exhaustion is adopted and an “insurer-only” payment requirement is enforced, the insured’s coverage losses are limited to the policy period in which the particular underlying policy is situated. For

coverages, such as the involved in most cases where the “continuous trigger” “all sums stacking” approach is applied. See text and notes 40-46, where the issue of horizontal exhaustion in the duty to defend context is discussed.

³⁸ *Ali v. Federal Ins. Co.*, 719 F. 3d 83 (2d Cir. 2013).

³⁹ See, e.g., *Qualcomm, Inc. v. Certain Underwriters at Lloyd’s, London*, 73 Cal. Rptr. 3d 770 (4 Cir. 2008). The policy provided: “Underwriters shall be liable *only after* the insurers under each of the Underlying policies *have paid or have been held liable to pay the full amount* of the Underlying Limit of Liability.” *Id.* at 778. The court held that the policy language did not permit the insured to make a bridging payment to allow access to the excess insurance coverage. *Id.* at 774-75.

⁴⁰ See generally Jeffrey W. Stemple, *An Analytical “Gap”: The Perils of Relentless Enforcement of Payment-by-Underlying-Insurer-Only Language in Excess Insurance Policies*, 52 TORT TRIAL & INS. PRAC. L.J. 807, 863 (2017) (criticizing judicial willingness to enforce “insurer-only” payment language as unfair to insureds and as unduly protective of insurers).

example, in Schematic 2, if the insured settled a coverage dispute with Insurer A for less than policy limits, the immediate consequences would be limited to year 1 coverages if only the coverages in place for that policy year are in play. If horizontal exhaustion is required, the consequences of the settlement may now spread into years 2 through 6. The insurers providing overlaying coverage in years 2 through 6 can now claim that their coverages are not triggered because *all* underlying insurance has not been exhausted by the payment of policy limits. Thus, not only does the tower of coverage in the specific policy period collapse, but application of horizontal exhaustion results in the entire insurance program, covering all the triggered policies, collapsing.

There is no substantial reason to permit the insurers in years 2 through 6 to withhold coverage based upon the resolution of a year 1 coverage dispute. Insurers B through D may be seen to have contracted that their obligations would be conditioned on the full performance by Insurer A of its obligations. That concession cannot be comfortably extended to the years 2 through 6 insurers except through a very generous reading of “Other Insurance” or similar provisions that may generally reference other underlying policies. Allowing the year 2 through 6 insurers to claim the settlement as a defense to payment by them renders the “uber” policy illusory. Similarly, horizontal exhaustion allows insurers B through D to defer their coverage obligations until exhaustion of underlying coverage in years 2 through 6, coverage that was not in existence at the time insurers B through D placed their coverage. There is no reason why Insurers B through D should be allowed to escape their coverage obligations based on coverages obtained by their insured after the B through D policies were obtained. Surely such actions by the insured are completely fortuitous to the decision by insurers B through D to provide coverage to the insured. Any settlement by an insured of a coverage dispute with one insurer would require the unanimous consent of all insurers whose policies have been triggered by the insured event to not treat the settlement as affecting other coverages. And because refusal to consent would preserve a complete coverage defense, every overlaying insurer would be incentivized to withhold consent.

Horizontal exclusion would in this context provide a windfall to insurers that cannot be justified by any underwriting or policy

considerations.⁴¹ A good faith settlement between the insured and an insurer promotes the efficient resolution of disputes and is consistently recognized as a desirable goal of civil adjudication. Therefore, to the extent courts would recognize and enforce “insurer-only” payment requirements, courts should not impose a horizontal exhaustion requirement, particularly in the continuous trigger, successive coverage context where doing so would allow *all* overlaying insurers to escape coverage obligations based on a single failure to pay policy limits by any single underlying insurer.

One consideration might, however, militate in favor of allowing overlaying insurers to withhold coverage when an underlying insurer has failed to completely exhaust its policy limits. This occurs when the underlying insurer has defense obligations, the cost of which are not credited against policy limits.⁴² In such a case, overlaying insurer may claim they have bargained for two layers of protection, (1) an indemnity protection based on policy limits, and (2) an unlimited defense obligation until the policy limits have been exhausted.

In the usual case, an overlaying insurer reasonably expects that defense costs will be borne by the underlying insurer until the underlying

⁴¹ See *Carter-Wallace, Inc. v. Admiral Ins. Co.*, 718 A.2d 1116, 1123 (N.J. 1998) (rejecting insurer proposed horizontal exhaustion in continuous trigger, successive coverage context).

⁴² In liability coverages defense costs are traditionally paid pursuant to the Supplemental Payments provision of the policy and are not charged against the policy’s indemnity limits. See KEETON, WIDISS & FISCHER, *INSURANCE LAW* § 8.1(e), at 790 (2d ed. 2016). The major exception to this principle occurs with respect to professional liability coverages where all or a portion of the insurer’s cost of defending the insured may be offset against the insurer’s indemnity obligation. *Id.* at n.51; see generally JERRY & RICHMOND, *UNDERSTANDING INSURANCE LAW* § 111[K] (4th ed. 2017). Defense costs can be substantial. See, e.g., *Biomass One, L.P., v. Imperial Cas. & Indem. Co.*, 968 F.2d 1220 (9th Cir. 1992) (Table) (noting that in defending policy’s indemnity limits of \$2 million, defense costs of \$1.9 million had been incurred); *Montgomery Ward & Co., Inc. v. Imperial Cas. & Indem. Co.*, 97 Cal. Rptr. 2d 44, 46-47 (2d Cir. 2000) (noting several instances where defense costs associated with the defense of specific contamination claims exceeded indemnity limits by a factor of approximately 2.5, e.g., policy limits of \$1 million and defense costs in excess of \$2 million; moreover, in each case the defense costs substantially exceeded the indemnity costs actually paid, in two of the instances by a factor of ten).

policy is exhausted by payment of judgments and settlements. A within limits settlement of a coverage dispute between the insured and an underlying insurer effectively denies the overlaying insurer this measure of protection. And, as many courts have noted, the presence of a defense obligation correlates with the price differential between underlying (primary, or working, layer of coverage) and overlaying coverages.⁴³ Allowing an underlying insurer to escape its defense obligations may be seen as unfair to the overlaying insurer, which has relied on the underlying policy’s defense obligation in pricing the overlaying policy. Simply allowing the insured to pay (or credit) the difference between the indemnity limits and the amount actually received from the underlying insurer does not fully make the overlaying insurer whole because that payment (or credit) does not reflect defense payments that would have been borne by the underlying insurer until the indemnity limits were, in fact, fully exhausted by payment of judgments or settlements. This consideration is limited to defense costs and does not extend to indemnity obligations that accrue as a result of settlement or judgment.

The proposed Restatement addresses this issue in a slightly different manner. It notes that the premium charged by overlaying insurers may be based on the expectation that underlying insurers will more competently evaluate and resolve claims using policy money than will insureds using their own money.⁴⁴ But this distinction between insurer and insured acumen amounts to little if does not encompass the defense obligation, which is generally unlimited and not tied to the policy limits. A rational overlaying insurer may reasonably expect the great majority of claims to be resolved within the policy limits as long as defense costs are excluded from the

⁴³ See, e.g., *Ali v. Fed. Ins. Co.*, 719 F.3d 83, 91 (2d Cir. 2013) (quoting *Gabarick v. Laurin Mar. (Am.), Inc.*, 649 F.3d 417, 422 (5th Cir. 2011)); See also *Zurich Ins. Co. v. The Heil Co.*, 815 F.2d 1122, 1126 (7th Cir. 1987); See also *Maricopa Cty. v. Fed. Ins. Co.*, 157 Ariz. 308, 310 (Ariz. Ct. App. 1988).

⁴⁴ RESTATEMENT OF THE LAW OF LIA. INS. §42 cmt. d (AM. LAW INST., Tentative Draft No. 1, 2016). The reasoning behind the Restatement’s position is debatable. One would think insureds would be as careful with their money as insurers are with their money in the contexts where the “continuous trigger” is most often used – mass torts implicating large, sophisticated insureds who often have risk professionals and platoons of lawyers to advise them.

calculation. This expectation is reflected in the cost differential between primary and excess insurance⁴⁵ and the value consistently recognized by courts and commentators that the insurer's defense obligation provides insureds.⁴⁶

While the costs of the defense clearly affect the pricing of primary and excess coverages, it would be a mistake to place too much emphasis on that fact in the "continuous trigger" "all sums stacking" context. When coverage involves only a single policy period (concurrent coverage), an overlaying excess insurer has a legitimate interest in the assumption of defenses cost by the underlying insurer providing primary coverage. Whether that duty is discharged by the primary insurer or the insured should be irrelevant, unless the overlaying insurer can demonstrate actual prejudice

⁴⁵ For example, in *Schmitz v. Great Am. Assur. Co.*, 337 S.W.3d 700 (Mo. 2011), the court noted that the primary policy (with \$1 million policy limits) cost \$8,386, while the excess policy (with \$4 million policy limits) cost \$4,000.

⁴⁶ See, e.g., *Montrose Chem. Corp. v. Superior Court*, 6 Cal. 4th 287, 295-96 (Cal. 1993) ("The insured's desire to secure the right to call on the insurer's superior resources for the defense of third party claims is, in all likelihood, typically as significant a motive for the purchase of insurance as is the wish to obtain indemnity for possible liability. As a consequence, California courts have been consistently solicitous of insureds' expectations on this score."); Eileen B Eglin & Stephen D. Straus, *Classifying RI/FS Costs under a Policy of Comprehensive General Liability Insurance or Defense*, 5 FORDHAM ENVTL. L. REV., 385 at 387 (2011) ("Depending upon the policy language, defense costs will either count towards the stated coverage limit of the policy or they will be exclusive of the limit. Under a cost-exclusive policy, the insurer's coverage obligation has the potential to be far greater than the stated indemnity limit. This is because defense costs in a cost-exclusive policy will not serve to impair the liability limit; only payments for damages in the form of judgments or settlements impair or exhaust the limits of a cost-exclusive policy. Defense costs under a cost-exclusive CGL policy can eclipse the stated policy limit where no settlements or judgements equaling the limits are sustained."); See also KENNETH S. ABRAHAM & DANIEL SCHWARCZ, *INSURANCE LAW AND REGULATION* 577 (6th ed. 2015) ("Most primary liability insurance policies not only provide indemnity to the insured; they also provide the right to a defense of all claims alleging liability that would be covered by the policy if the allegations were true. This coverage provides important 'litigation insurance,' since the costs of defending against even unsuccessful lawsuits can be substantial.")

if the defense is maintained by the primary insurer or the insured. In the “continuous trigger” context, however, the loss of a primary policy with a defense obligation, does not affect the obligation of other primary insurers to provide a defense. Courts has consistently recognized the primacy of the duty to defend in this context; therefore, any and all insurers in the “continuous trigger” context, whose policies contain a defense obligation, must provide a defense if their policy is triggered. Thus, failure of a primary insurer to provide a defense is unlikely to have an impact on an overlaying insurer because a complete defense will be provided by the other primary insurers. Permitting an overlying insurer to escape its indemnity obligations because one or even several underlying insurers failed to provide a defense would be a complete exaltation of form over substance in the “continuous trigger” context.

Some courts have suggested that concern over collusion between an insured and an underlying insurer is an independent justification for requiring actual payment of indemnity limits by an insurer to exhaust the policy. For example, in *Ali v. Federal Insurance Company* the court distinguished *Zeig v. Massachusetts Bonding & Insurance Company* on the ground that *Zeig* involved a property insurance loss that became fixed when it occurred. *Ali*, on the other hand, involved liability insurance coverage (Directors & Officers policy), which involved the insureds “obligations to pay third parties.” The *Ali* court agreed with the lower court that this obligation to pay could incentivize the insured and its underlying insurers to structure settlements manipulatively to move payment away from the underlying insurers and to the overlaying insurers.⁴⁷

As a speculative proposition, it is, of course, possible that an insured could collude with one of its insurers to the prejudice of another of its insurers. Instances of this occurring have been reported,⁴⁸ and, if it occurs it seemingly would more likely occur between an insured and an insurer the insured had a working relationship with, e.g., the primary layer insurer, and to the prejudice of an insurer the insured had only a distant relationship with, e.g., an excess insurer somewhere in the tower of coverage provided by the

⁴⁷ *Ali*, 719 F.3d at 93-94.

⁴⁸ See e.g., *Kaiser Found. Hosps. v. North Star Reinsurance Corp.*, 153 Cal. Rptr. 678, 682 (Dist. Ct. App. 1979) (involving collusion between insured and primary insurer to allocate dates of loss to transfer loss exposure away from primary insurer and onto excess insurer).

insurance coverage program. That said, it is difficult to conceive of factual scenarios involving collusion that do not involve allowing an underlying insurer to escape its defense obligation. If only indemnity is involved, allowing the underlying insurer to buy its way out of coverage does not financially impact an overlaying insurer because to preserve the overlaying insurance the insured with have to assume the discharged insurer's indemnity obligation in some manner, e.g., by payment (or credit) up to the underlying limits. Thus, a less than limits payment by an underlying insurer has no financial consequences to an overlaying insurer insofar as indemnity is concerned because the overlaying insurer will receive a credit against the loss equal to any difference between the policy limits and the insured-insurer coverage settlement. For example, using Schematic 1, assume Insurer Alpha, with indemnity limits of \$1 million, settles a coverage dispute with the insured for a payment of \$500,000. The claimant and the insured settle a dispute (which only affects the year 1 policies) for \$2 million. Assume further, the Epsilon policy has \$3 million policy limits. Epsilon would pay \$1 million. This reflects the full value of underlying insurance (\$1 million) being credited against the settlement.

This suggests that concern over collusion and manipulation of settlements is misguided. Aside from avoidance of defense costs, insureds and underlying insurers have little or no reason to collude because they are not actually exporting any of their costs to overlaying insurers. Protecting an overlaying insurers' reliance on underlying insurers absorbing defense costs until their policy limits are exhausted by payment of judgments or settlements is not a standalone justification for not treating a within limits settlement between an insured and insurer as exhausting the limits. As noted earlier, the defense will be provided by other primary insurers or the insured so that the overlaying insurer is not prejudiced. A focus on collusion or manipulation adds nothing to the analysis whether the overlaying insurance should be deemed to be "triggered" with respect to excess coverage obligations.

III. EXHAUSTION OF COVERAGE THROUGH NON-CUMMULATION/PRIOR INSURANCE PROVISIONS

Non-Cumulation/Prior Insurance provisions refer to insurance policy terms that provide that the applicable policy limits may be offset by insurance under another policy available to the insured. Non-Cumulation/Prior Insurance provisions are similar to "Other Insurance" provisions in that both seek to reduce the insurer's obligations when the insured is also an insured under other insurance policies that cover the same

insured event. Many courts, however, limit the application of “Other Insurance” provisions to insurers on the same risk in the same policy period and refuse to apply the provision in cases of successive coverage.⁴⁹ Non-Cumulation/Prior Insurance provisions are said to address (from the insurer’s vantage point) the problem raised by the “continuous trigger” “all sums stacking” approach.⁵⁰ Non-Cumulation/Prior Insurance provisions allow the insurer to use payments owed or made under prior policies that apply to the same risk under a continuous trigger to reduce obligations owed under other policies made applicable to the risk by the same continuous trigger. For example, looking at Schematic 3, on page 286, assume under “continuous trigger” “all sums stacking” Insurer Beta’s year 1, year 2, and year 4 policies have all been triggered. A Non-Cumulation/Prior Insurance provision is intended by the insurer (here Beta) to allow it to offset payments made by it from policies at the same layer of coverage. For example, a payment of \$1 million on the year 1, 1st layer excess Beta policy would be credited against the year 4, 1st layer Beta policy. Thus, instead of Beta paying \$1 million on each policy, or \$2 million total, Beta would pay only \$1 million on both policies. The result would be to negate “all sums stacking.” Courts have enforced Non-Cumulation/Prior Insurance provisions when the payments are made by the same insurer⁵¹ and when made by different insurers.⁵² On the other hand, a number of courts have refused to enforce Non-

⁴⁹ See *Ohio Cas. Ins. Co. v. Unigard Ins. Co.*, 564 F.3d 1192, 1196-97 (10th Cir. 2009) (collecting conflicting authorities). The court certified the question to the Utah Supreme Court which held that “Other Insurance” provisions did not apply to successive insurers on the same risk. *Ohio Cas. Ins. Co. v. Unigard Ins. Co.*, 268 P.3d 180, 183 (Utah 2012); see also *In re Viking Pump*, 52 N.E.3d 1144, 1157 (N.Y. 2016) (same).

⁵⁰ See *Viking Pump*, 52 N.E.3d at 1152. The court noted that Non-Cumulation/Prior Insurance provisions are not applicable when a jurisdiction adopts the “pro-rata” rather than the “all sums with stacking” method of allocation because under “pro rata” allocation the insurer is only liable under each triggered policy for that portion of the total loss that occurred in the particular policy’s period of coverage. *Id.* at 1153.

⁵¹ *Olin Corp. v. Am. Home Assurance Co.*, 704 F.3d 89, 102-104 (2d Cir. 2012) (applying New York law) (*Olin III*).

⁵² *Olin Corp. v. OneBeacon Am. Ins. Co.*, 864 F.3d 130, 151 (2d Cir. 2017) (applying New York law) (*Olin IV*).

Cumulation/Prior Insurance provisions, usually on the ground that the provision is ambiguous.⁵³

Schematic 3

\$4 million	Tau		Tau	Omega
\$3 million		Gamma		
\$2 million	Beta			Beta
\$1 million				
	Year 1 Policies	Year 2 Policies	Year 3 Policies	Year 4 Policies

Although Non-Cumulation/Prior Insurance provisions, when enforced, may have their greatest impact on the selection of a continuous trigger allocation method,⁵⁴ the provisions also may influence the exhaustion method adopted by the court. The adoption of the “all sums stacking” approach, because of the presence of a Non-Cumulation/Prior Insurance provision, encourages the adoption of vertical exhaustion over horizontal exhaustion.⁵⁵ If the excess insurance policy expressly provides that its limits are reduced by other payments, those other payments determine whether the excess insurance is available. Requiring exhaustion of all underlying insurance simply adds an additional requirement that is inconsistent with the explicit requirements of the Non-Cumulation/Prior Insurance provision because exhaustion here simply duplicates what is accomplished by an

⁵³ A.B.S. Clothing Collection, Inc. v. Home Ins. Co., 41 Cal. Rptr. 2d 166, 170-173 (Cal. Ct. App. 1995).

⁵⁴ As noted previously, many courts find that the presence of a non-cumulation/prior-insurance provision in a policy is inconsistent with an insurer’s claim that “pro rata” allocation should be adopted. As noted in *In re Viking Pump*:

[P]olicies containing non-cumulation clauses or non-cumulation and prior insurance provisions...all sums is the appropriate allocation method...[I]t would be inconsistent with the language of the non-cumulation causes to use pro rata allocation here.

52 N.E.3d 1144, 1153 (N.Y. 2016).

⁵⁵ *Id.* at 1156 (stating that “vertical exhaustion is conceptually consistent with an all sums allocation”).

enforceable Non Cumulation/Prior Insurance provision. If, on the other hand, horizontal exhaustion is applied, a question may arise whether a policy that is not fully paid, because of credits applied from other policy payouts, has been exhausted. Overlaying insurers will likely argue that a credit does not constitute “actual payment” of policy limits by payment of judgments or settlements. The presence of Non- Cumulation /Prior Insurance provisions, thus, will likely complicate accessing the insured’s total coverage program. And, if the credit is not treated as an “actual payment,” application of a Non-Cumulation /Prior Insurance provision may possibly foreclose recovery from the overlaying policies due to non-exhaustion of the underlying policy.

IV. INSOLVENCY OF UNDERLYING INSURERS

Insolvency has been addressed most commonly in “continuous trigger” coverage contexts in the area of allocation.⁵⁶ It also arises in the more traditional context of the overlaying insurer’s “drop down” obligations, which typically involves a single policy period containing a single tower of coverage, i.e., concurrent coverage.⁵⁷ There has been relatively little judicial

⁵⁶ See John T. Waldron & Sara N. Brown, *New Jersey Allocation Law: How to Calculate Solvent and Insolvent Insurers Pro-Rata Shares for Long-Tail Claims*, 50 TORT TRIAL & INS. L. J. 747 (2015).

⁵⁷ See, e.g., *Mission Nat’l Ins. Co. v. Duke Transp. Co.*, 792 F.2d 550 (5th Cir. 1986) (holding that insolvency of underlying insurer did not trigger obligation on the part of overlaying insurer to assume the insured’s defense). This issue has generated substantial judicial disagreement in older decisions when there was more variation in coverage language. Cf. *Reserve Ins. Co. v. Pisciotta*, 640 P.2d 764 (Cal. 1982) (holding that policy language triggering overlaying insurers duty to assume the insured’s defense, when “amount recoverable” under the underlying policy was paid, was ambiguous and could be reasonably understood to trigger overlaying insurer’s obligation when insolvent underlying insurer paid all that it could, even though that amount was less than policy limits); *with Moorpark Indus. Inc. v. W. Emps. Ins. Co.*, 429 N.W.2d 213, 218 (Mich. Ct. App. 1988) (stating that term “amount reasonable” is not ambiguous); see generally Jane M. Draper, Annotation, *Primary Insurer’s Insolvency as Affecting Excess Insurer’s Liability*, 85 A.L.R.4th 729, 757-63 (1991) (collecting conflicting decisions whether term “amount recoverable” is ambiguous). See *id.* at 763-75. Modern policies do not use the phrase “amount recoverable” to trigger

discussion of the effect of insolvency on the obligations of overlaying insurers when a “continuous trigger” is applied with the result that coverage is successive rather than just concurrent. That is somewhat surprising given the fact that insurer insolvency will likely be encountered in the “continuous trigger” context given the financial demands continuous trigger theories place on insurers, particularly when “all sums stacking” is imposed.⁵⁸

Traditionally, an indemnity obligation required the payment of the debt for the indemnity to be triggered. Thus, if B was indebted to A and C agreed to indemnify B, A’s ability to compel C to perform required that B sustain a loss. If B was insolvent, C could escape because B sustained no loss, being insolvent. Such a rule would be beneficial to insurers, who traditionally occupy the role of C, while their insureds occupy the role of B; however, most states and standard form policy language have taken this defense away.⁵⁹ Those statutes and policy language apply, however, only when the Bs (insureds) of the world become insolvent. What happens when the Cs of the world (insurers) become insolvent?

Standard form primary policy language does not directly address insurer insolvency; therefore, the issue is usually addressed by examining the policy language that activates the overlaying coverage. In this sense the issue of insurer insolvency is similar to the issues raised by settlements between the insured and an insurer for less than policy limits discussed in Part II of this paper. Does the settlement (here, insurer insolvency) satisfy a requirement in the overlaying policy that conditions coverage on (1) the

coverage; rather, they use the phrased “reduced or exhausted by payment,” or a variant of the phrase, which is generally interpreted to require exhaustion by payment of the policy’s actual, stated limits. *See id.* at 746-57. The term “collectible,” which some courts have found to be “ambiguous” is still, however, used in many excess policies.

⁵⁸ I could not find any data supporting or disproving this assumption. There is no disagreement, however, that adoption of continuous triggers has significantly increased insurer financial obligations to insureds. If it didn’t insureds would not press for it and insurers would not fight tooth and nail against it! Adam Raphael’s book *Ultimate Risk* does attribute Lloyds financial crisis in the 1990’s in significant part to mass tort claims that, through the continuous trigger theory, allowed insureds to aggregate coverage for particular losses across multiple policy periods.

⁵⁹ *See* David Gray Carson, *Indemnity, Liability, Insolvency*, 25 CARDOZO L. REV. 1951 (2004) (discussing distinction between indemnity and liability obligations when the obligor becomes insolvent).

exhaustion of limits by payment of judgments or settlements or (2) for “ultimate net loss” in excess of underlying scheduled or other insurance?

Courts have split as to whether insurer insolvency will result in the inability to satisfy an “ultimate net loss” provision. Much turns on the specific language of the provision and the willingness of a court to deem the provision clear or ambiguous. For example, in *Reserve Insurance Company v. Piscotta*, the California Supreme Court held that “ultimate net loss” language that was tied to the “amount recoverable” was not expressly tied to policy limits; therefore, the insolvency of an underlying insurer required that the overlaying insurer “drop down” and provide coverage.⁶⁰ Not all courts, however, agree with this construction of the phrase “amount recoverable” as ambiguous.⁶¹

In most instances, insurers use language in the “ultimate net loss” clause that courts deem reasonably informs the insured that only full payment of the underlying limits will allow the insured to access the overlaying insurance; consequently, the insolvency of an underlying insurer is often a risk that is borne by the insured, not the overlaying insurer(s). For example, in *Mission National Insurance Company v. Duke Transportation Company*, the overlaying insurer (Mission) conditioned its coverage “to the ultimate net loss the excess of...the limits of the underlying insurance as set out in the attached schedule...”⁶² As a result of the insolvency of the underlying insurer (Northwest), that attachment threshold could not be met. The insured (Duke) argued that this event (insolvency) meant that Mission’s obligations became immediate. In other words, because Northwest could no longer perform its obligations, Northwest’s obligations should be deemed to have been performed. The court rejected this argument holding that Northwest’s insolvency did not excuse the policy requirement that losses as specified be paid before Mission’s obligations would be activated.⁶³

Thus, initially at least, “insolvent insurers cases” tract “settlement within coverage limits” cases. In both cases, the critical issue is whether the

⁶⁰ *Pisciotta*, 640 P.2d 764, discussed in note 57. In effect, one reasonable interpretation of the “amount recoverable” language was that it referred to the amount actually recoverable (or recovered) from the insolvent insurer, not the amount set by the insolvent insurer’s policy limits.

⁶¹ See cases cited *supra* note 57. Nonetheless, the phrase “amount recovery” is no longer used in excess insurance policies.

⁶² *Mission Nat’l Ins. Co.*, 792 F.2d 550, discussed in note 57.

⁶³ *Id.* at 553.

requirement that the underlying limits have been paid in full. In both cases, many courts give primacy to policy language to determine whether full, actual payment of policy limits by the underlying insurer is required as a condition precedent to reach the overlaying insurance. In both cases, modern policy language is often read by courts today as requiring payment of the full limits by the insurer before overlaying insurance may be accessed.

In one way, however, insurer insolvency cases differ from “within limits settlement” between insureds and insurers—the insolvent insurer is liquidated in an administrative proceeding and confirmed by judicial process and review. Unlike the private bargain between the insured and the underlying insurer, insolvent insurer proceedings are specifically designed to provide a fair and efficient resolution of claims given the resources available. The presence of judicial oversight and review may persuade a court that sufficient protections exist so that the interests of overlaying insurers are protected, such that once the claim(s) is/are resolved the court may deem the attachment point satisfied. Judicial approval of a plan of liquidation of the insolvent insurer may be treated as a judicial determination that the attachment point was reached, even though the actual amount provided is less than the policy limits.⁶⁴

V. LOSS ALLOCATION AGREEMENTS BETWEEN INSURED

Insureds often enter into loss allocation agreements with each other and the meshing of these agreements with the risk transfer provided by insurance has proved difficult.⁶⁵ For example, an insured (Contractor) may

⁶⁴ See *Canon Elec., Inc. v. Ace Prop. & Cas. Co.*, Case No. BC 290354 (Superior Court, County of Los Angeles, August 17, 2017) (copy on file with the author). In this matter the trial court concluded that a judicial determination in liquidation of the amount the insolvent insurer would pay satisfies the exhaustion requirement even though the amount paid is less than policy limits. *Id.* at 50-55 (applying New Hampshire law). In effect the court concluded that a court order that the insolvent insurer had paid its obligations under the policy was equivalent to the actual payment of policy limits.

⁶⁵ See 4 PHILIP L. BRUNER & PATRICK J. O’CONNOR, JR., BRUNER & O’CONNOR CONSTRUCTION LAW § 11:554 (2d ed. 2017) (noting common use of risk-transfer agreements in construction industry and uncertainty, in the absence of specific identification of the problem in the insurance policy, regarding the primary of the risk-transfer agreement on the policy); Jeremiah M Welch & Julian D. Ehrlich, *Horizontal Exhaustion: Challenges and*

enter into an agreement with a third person (Subcontractor) in which Subcontractor agrees to exculpate the Contractor for all losses arising from Subcontractor’s performance of the agreement.⁶⁶ Contemporaneously, the parties will agree that some or all of the parties will become additional insureds on a party’s existing coverage, to which that party’s insurer(s) agree. To what extent, if at all, should the indemnity agreement affect insurance coverage available to the parties to the indemnity agreement, here, Contractor and Subcontractor? And to the point of this paper, to what extent, if all, should the indemnity agreement affect exhaustion requirements regarding overlaying levels of insurance coverage provided by excess insurers?

An example may help illustrate the problem. Subcontractor agrees to perform work on a construction project being built by Contractor. As part of the agreement, Subcontractor agrees to name Contractor as an additional insured on its (Subcontractor’s) primary liability insurance policy. Contractor has primary liability coverage with Alpha Insurance Company and excess coverage with Beta Insurance Company. Subcontractor has primary liability coverage with Omega Insurance Company and excess coverage with Theta Insurance Company. Both excess coverages purport to

Solutions, 40 NYSBA TORTS, INS. & COMPENSATION L. SEC. J., Summer 2011 at 20.

⁶⁶ This is frequently accomplished by a “waiver of subrogation” rights which bar direct liability to the extent there is insurance in place that covers the loss. The effect of such a provision is to prevent the insurer that pays the loss from claiming reimbursement. These waivers are generally enforced and may not be used by insurers to refute coverage. *See* BRUNER & O’CONNOR, JR., *supra* note 65; *see also* 1 SCOTT C. TURNER, *INSURANCE COVERAGE OF CONSTRUCTION DISPUTES*, §5:7 (2d ed. 1999). A pre-loss release of liability has been treated the same as a waiver of subrogation rights. *See* *Great N. Oil Co. v. St. Paul Fire & Marine Ins. Co.*, 189 N.W.2d 404 (Minn. 1971). However, this decision involved a first party property insurance policy. In the field of construction disputes many jurisdictions restrict efforts to shift liability beyond that accomplished by insurance by barring insurers from obtaining reimbursement against those deemed additional insureds. *See* BRUNER & O’CONNOR, JR., *supra* note 65; *see also* Jay M. Zitter, Annotation, *Insurance: Subrogation of Insurer Compensating Owner or Contractor for Loss under “Builder’s Risk” Policy Against Alleged Negligent Contractor or Subcontractor*, 22 A.L.R.4th 701 (1983).

be excess to the scheduled underlying primary policies and “all other collectible underlying policies.” Finally, Subcontractor has entered into a separate agreement with Contractor in which Subcontractor agrees to indemnify Contractors for any losses resulting from Subcontractor’s performance of its work on the job site pursuant to its contract with Contractor.

Worker is injured on the job site and sues both Subcontractor and Contractor. The claim exceeds the limits of both primary policies, but not the limits of either excess policy. Both Subcontractor and Contractor tender Worker’s claim to all the insurers. Does the presence of the indemnity agreement affect the obligation of Beta Insurance Company (Contractor’s excess) or Theta Insurance Company (Subcontractor’s excess)? Does the indemnity agreement between the insureds affect the determination whether the underlying insurance is “collectible”?

Both excess insurers could argue that under principles of horizontal exhaustion both the Alpha policy (Contractor’s primary) and the Omega Policy (Subcontractor’s primary) must exhaust before either excess insurer must step forward. Does the indemnity agreement between Contractor and Subcontractor change that by shifting responsibility for the loss away from Contractor (and Alpha)?

More importantly, can Theta (Subcontractor’s excess) now disclaim liability on the ground that the Alpha policy has not been exhausted and, therefore, under principles of horizontal exhaustion, it’s (Theta’s) policy has not been triggered? In other words, Theta would contend that the Alpha policy remains “collectible” insofar as the Theta policy is concerned, even though by operation of the insured’s indemnity agreement the loss will be borne by Subcontractor (Theta’s insured)⁶⁷

Most decisions addressing whether insured–insured loss allocation agreements, such as the indemnity agreement between Contractor and Subcontractor, affect insurer obligations have involved disputes between primary insurers. In this context, courts have divided whether one insurer may use a loss allocation agreement, to which it (the insurer) is not a party,

⁶⁷ I am assuming here that Contractor’s Non-liability is the expected outcome. Unlike the insurer’s defense obligation which can be triggered by potential liability within coverage, the insurer’s indemnity obligation rests on actual liability within coverage. However, as a practical matter, the insurer cannot know whether the insured will be deemed liable on nonliable. In the course of defending the insured, the insurer may be forced to make difficult decisions regarding settlement when its coverage obligations are uncertain because the liability of its insured is uncertain.

to escape its obligations under its insurance policy. One line of decisions holds that the insurer may rely on the insured’s explicit loss allocation decisions; another line of decisions holds the insurer to its commitments in its insurance policy and refuses to allow the insurer to use agreements to which it is not a party to escape obligations it has contractually assumed.⁶⁸ For the most part, however, these decisions do not address the issue of exhaustion because they do not involve excess insurance.

Several decisions, however, have specifically involved disputes between primary and excess insurers where one of the insurers is using a loss allocation agreement, to which the insurer it is not a party, to trigger or avoid liability.

In *Bovis Lend Lease LMB, Inc. v. Great American Insurance Company*,⁶⁹ the Contractor and Subcontractor had separate policies and the Subcontractor agreed to add the Contractor as an additional insured to its (Subcontractor’s) policies “without contribution by the Contractor’s own insurance.” A claim was made that would trigger the attachment point of the Subcontractor’s excess insurance, so a dispute arose as to the priority of involved policies. The *Bovis* court held that in the absence of policy language adopting or permitting the adopting of exculpatory agreements between insureds allocating responsibility for a loss, priority of coverage would be determined by policy language and judicially developed coverage rules, here horizontal exhaustion. The court concluded that excess policies would not be triggered until all the underlying primary policies had been exhausted.⁷⁰

In another New York case, however, the court appeared to recognize a loss allocation agreement as excusing the requirement that an underlying policy be horizontally exhausted before the attachment point of an overlaying excess policy would be triggered. In *Indemnity Insurance Company of N.A. v. St. Paul Mercury Insurance Company*,⁷¹ insurers of the

⁶⁸ See generally SCOTT M. SEAMAN & JASON R. SCHULZE, ALLOCATION OF LOSSES IN COMPLEX COVERAGE CLAIMS § 5.4[f] (2018); BRUNNER & O’CONNOR, JR., *supra* note 65.

⁶⁹ See *Bovis Lend Lease LMB, Inc. v. Great American Insurance Company*, 855 N.Y.S. 2d 459, 462 (App. Div. 2008).

⁷⁰ *Bovis Lend Lease v. Great Am. Ins. Co.*, 855 N.Y.S.2d at 464, 470-71.

⁷¹ *Indem. Ins. Co. v. St. Paul Mercury Ins. Co.* 900 N.Y.S.2d 24 (N.Y. App. Div. 2010). Both decisions were from the 1st Department, but the

indemnitor (Subcontractor) accepted tenders of the indemnitee's (Contractor's) defense from the indemnitee's insurer (St. Paul). The indemnitor's insurers, (Royal (primary) and Indemnity Insurance Company of N.A. (excess)), subsequently settled the claim, with Royal paying its limits (\$1 million) and Indemnity Insurance Company paying \$2 million. Indemnity Insurance Company, in turn, sought reimbursement from the indemnitee's insurer, St. Paul. The court rejected Indemnity's claim that St. Paul's obligations were antecedent to its own and required exhaustion. Under the loss allocation agreement, the obligations of the indemnitee (Contractor) and its insurer (St. Paul) would "pass through" to the indemnitor (Subcontractor) and its insurer (Indemnity).

This conflict in approaches is further complicated by the fact that the decided cases involve concurrent coverages rather than successive coverages applicable to a loss under a "continuous trigger" approach. Again, this illustrates the consequences of adopting a horizontal exhaustion requirement that assumes an even, rising tide progressively exhausting layers of coverage. In reality that rising tide must confront numerous obstacles that may interfere with the smooth upwards movement that the rising tide metaphor assumes.

Courts generally have not addressed the issue of the effect of a risk-transfer agreement between insureds on the coverage obligations of insurers in terms of exhaustion. But the effect of choosing to give, or not give, primacy to the risk-transfer agreement may affect how coverage is sequenced. If, on the one hand, the court gives primacy to the insureds' risk-transfer agreement, coverage is limited to the party that has assumed the risks. If vertical exhaustion is used, the obligation of overlaying insurers is determined by looking at coverage within the specific underlying policy period that is activated. However, if horizontal exhaustion is applied, giving primacy to the loss-allocation agreement may result in a loss of coverage because one or more insurance coverages in the insured's insurance program will not be exhausted. If, on the other hand, the court does not give primacy to the insured's risk-transfer agreement, the choice between horizontal and vertical exhaustion will be determined by the coverage language in each insurance policy. As noted in Part I of this paper, the literal terms of most insurance policies can be read, and have been read by courts, as providing for horizontal exhaustion as the pathway to accessing overlaying coverage. In this case, courts must determine whether horizontal exhaustion is the proper method for accessing overlaying coverages when a "continuous

Indemnity Insurance Company court did not cite nor discuss *Bovis Lend Lease*.

trigger” “all sum stacking” approach (“the giant uber-policy”) is adopted. This point is addressed next.

VI. SUCCESSIVE COVERAGE AND EXHAUSTION

As noted in this paper, the issue of exhaustion has normally been considered in the context of concurrent coverage, that is layers of coverage within a single policy year. A few courts have considered the exhaustion issue in the context of successive coverage (caused by application of a “continuous trigger”), but the courts, for the most part, have not viewed concurrent coverage cases different from successive coverage disputes insofar as the exhaustion requirement is concerned. That, in this author’s opinion, is a mistake.

Proponents of horizontal exhaustion often use the metaphor of a rising tide, successively exhausting coverage layer by layer. When exhaustion doctrine is looked at broadly, one sees many situations where underlying policies will not be fully exhausted, due to coverage issues specific to those policies. While the rising tide metaphor makes some sense when applied to layers of coverage within a single policy period, the metaphor is ill-suited to the situation presented in “continuous trigger” “all sum stacking” cases where multiple policy periods are involved and the insured’s coverage profile often differs substantially from year-to-year.

As a practical matter, when multiple policy periods are in play due to the application of a continuous trigger, it is highly likely that one or more of the underlying policies will not pay out its full limits due to one or more of the reasons set out in this paper. If horizontal exhaustion is applied, that failure to pay the full limits will, in effect, block the rising tide of exhaustion. Rather than a rising tide, a more appropriate way of seeing horizontal exhaustion is to envision a rising level of water that must proceed through a plethora of bottlenecks where upward movement will be blocked forever once the bottleneck of a single “unpaid in full” policy is encountered. That consequence is simply inconsistent with the “Uber” policy approach the California Supreme Court adopted in the State of California Continental Insurance Company decision. Horizontal exhaustion effectively transforms the “Uber” policy to a “Mini” policy by providing overlaying insurers with an escape card at the point any single underlying policy fails to pay its full limits. In effect, the “Uber” policy is an illusion if horizontal exhaustion is applied. More perniciously, horizontal exhaustion punishes the insured who acquires more insurance because more insurance increases the risk that one

of the additional insurance policies may fail to pay its policy limits, thus toppling what remains of the insured's coverage program. Or to paraphrase the Notorious B. I. G.: the more insurance a policy holder obtains, adoption of horizontal exhaustion means the more problems the insured has in collecting on any of the policies.⁷²

When successive coverage arises by operation of a "continuous trigger" "all sums stacking" approach, the question is squarely presented how the overlaying coverage should be accessed. Simply adopting a solution by arguing that horizontal exhaustion is more consistent with the nature of the relationship between primary and excess insurance is insufficient for several reasons. First, successive coverage simply presents a different context in which coverage questions must be resolved and ignores the fact that the "continuous trigger" doctrine was adopted to address problems caused by great societal problems (asbestos, environmental degradation, etc.) which could be addressed more efficiently and effectively through the cost sharing and cost spreading attributes of insurance. Second, horizontal and vertical exhaustion requirements are complicated doctrines. As shown in this paper, whether and when, if ever, an underlying policy is exhausted and whether and when, if ever, an overlaying policy must respond to a loss is often a question fraught with uncertainty in the concurrent coverage context; that question becomes exponentially more uncertain when coverage is expanded in the successive coverage context. Third, treatment of an exhaustion requirement in successive coverage cases should reflect the reasons and values that led to the recognition of successive coverage through adoption of the "continuous trigger" "all sum stacking" approach.

The approach that is plainly inconsistent with successive coverage is horizontal exhaustion. Adopting horizontal exhaustion, when coupled to the doctrine discussed in the paper, will often result in overlaying insurers completely escaping liability because one or more underlying policy in one or more years of the continuous trigger was not exhausted. The practical effect of adopting horizontal exhaustion is to limit successive coverage to primary insurers and give all overlaying insurers arguments to escape coverage all together. That would render the "continuous trigger" "all sums stacking" approach a toothless doctrine.

⁷² NOTORIOUS B.I.G., *MO MONEY, MO PROBLEMS* (Bad Boy Records and Arista Records 1997).

TRANSITIONING TO A NEW VIEW: COMING TO SEE HEALTH INSURANCE COVERAGE FOR GENDER DYSPHORIA IN A NEW LIGHT

ROBERT A. WHITNEY*

TABLE OF CONTENTS

INTRODUCTION	297
I. THE INITIAL REVIEW PROCESS	299
II. MENTAL HEALTH PARITY	305
III. UNFAIR INSURANCE PRACTICE	308
IV. SEXUAL ORIENTATION	313
V. SEX DISCRIMINATION	314
VI. TRANSITIONING TO A NEW VIEW	320
VII. THE AFTERMATH.....	322
CONCLUSION	326

INTRODUCTION

On June 20, 2014, the Massachusetts Division of Insurance (“the Division”) issued Bulletin 2014-03 (“the Bulletin”), entitled “Guidance Regarding Prohibited Discrimination on the Basis of Gender Identity or Gender Dysphoria Including Medically Necessary Transgender Surgery and Related Health Care Services.”¹ As set forth in the Bulletin, the Division concluded that the denial of coverage by health insurance companies for

* Robert A. Whitney is a consultant with over twenty-five years of experience in the insurance and reinsurance industry. He was Deputy Commissioner and General Counsel at the Massachusetts Division of Insurance from September 2011 through October 2015. All the views expressed herein are solely those of the author.

¹ COMMW. OF MASS, OFFICE OF AFF. & BUS. REG, BULLETIN 2014-03 (June 20, 2014), <https://www.mass.gov/files/documents/2017/11/21/bulletin%202014-03%20%28Gender%20Signed%29.pdf>.

gender transition-related medical care including gender assignment surgery, hormone replacement therapy, and other treatments based on an individual's gender identity or gender dysphoria was sex discrimination and prohibited under Massachusetts law.²

In issuing the Bulletin, the Division also concluded that the nearly uniform exclusion of coverage for gender identity or gender dysphoria-related treatment by Massachusetts health plans is considered prohibited sex discrimination because it would be a limitation on coverage based on the sex of the insured. As a result, the Division determined that any health care services that are ordinarily or exclusively available to individuals of one sex may not be denied based on the perceived gender identity of a person when the denial or limitation is due only to the fact that the insured is enrolled as belonging to the other sex, or has undergone, or is in the process of undergoing, gender transition.³

The Division also concluded that although a carrier may exclude coverage for a particular condition or treatment to the extent allowed by law, the insurer may not base such exclusion on gender identity or gender dysphoria. In this regard, the Division concluded that a carrier may not discriminate on the basis of an insured's or prospective insured's actual or perceived gender identity, sex stereotyping, or on the basis that the insured or prospective insured is a transgender person.⁴

On the same day that the Bulletin was issued, the administration of Governor Deval Patrick also announced that MassHealth, the Massachusetts Medicaid program, would cover gender re-assignment surgery as a standard benefit in its government health plan for lower-income persons and persons with disabilities.⁵ As reported by the Boston Globe at the time, the advocacy group Gay & Lesbian Advocates and Defenders ("GLAD") described these

² Gender dysphoria is the official diagnosis of individuals who view themselves as being different from their assigned birth sex. The term is often used to describe persons who experience significant dysphoria with respect to their gender identity, which is described as a feeling of acute hopelessness and discontentment with their own biological sex. See American Psychiatric Association, *Gender Dysphoria* (2013), <http://www.dsm5.org/Pages/Default.aspx>.

³ See *supra* note 1, at 1.

⁴ See *id.*

⁵ See Press Release, "Governor Announces Changes in Health Insurance Access for Transgender Community," MassEquality, June 20, 2014, <http://www.massequality.org/content/governor-announces-changes-health-insurance-access-transgender-community>.

two pronouncements as “historic” because at the time no other state had “announced in one fell swoop and this comprehensively, that medical care for transgender people is essential.”⁶

The determination by the Division that exclusions from health insurance coverage for gender transition-related medical care would no longer be permitted in Massachusetts was the culmination of an almost six-month review process by the Division where, at the time, I was the Deputy Commissioner and General Counsel. This Article explores how the Division reviewed the state of the law at the time, both on the federal and state level, to see if the strong prohibition in Massachusetts against discrimination under law also extended to prohibiting discrimination in healthcare coverage on the basis of gender identity or gender dysphoria.

I. THE INITIAL REVIEW PROCESS

In late 2013 and early 2014, advocacy groups such as GLAD and Health Law Advocates approached the Division asking it to declare that Massachusetts law precluded the exclusion of gender transition-related care from private insurance coverage, and that such exclusion was unlawful discrimination on the basis of gender identity or gender dysphoria. At the time, the majority of health insurers in Massachusetts that were subject to regulation by the Division excluded from their medical plans coverage of medical treatment for persons with gender dysphoria.⁷

In response to the requests from advocacy groups and individuals who were denied coverage under their Massachusetts health plans for gender transition-related medical care, including gender assignment surgery, hormone replacement therapy and other treatments, the Division began to

⁶ See Jeremy Fox, *Mass. to Cover Range of Transgender Medical Care*, BOSTON GLOBE (June 20, 2014), <https://www.bostonglobe.com/metro/2014/06/20/state-cover-gender-reassignment-surgery-and-hormone-treatment-for-transgender-patients/a9OPrvqdUPmRoiAQugVwEO/story.html>.

⁷ Gender dysphoria is the official diagnosis of individuals who view themselves as being different from their assigned birth sex. The term was often used to describe persons who experience significant dysphoria with respect to their gender identity, which is described as a feeling of acute hopelessness and discontentment with their own biological sex. See AM. PSYCHIATRIC ASS’N, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS, GENDER DYSPHORIA (5th ed. 2013).

review Massachusetts' own laws, as well as federal law and the law of other states, to determine whether health insurance carriers should be prohibited from excluding from coverage medical treatment related to gender dysphoria.⁸

As an initial matter, the Division looked to see whether there was any specific law in Massachusetts that would preclude such exclusions from being enforceable because of the insured's gender identity alone. For example, on November 23, 2011, Governor Deval Patrick signed into law Chapter 199 of the Acts of 2011, entitled "An Act Relative to Gender Identity" ("Chapter 199").⁹ This law added "gender identity" as a new protected characteristic under Massachusetts' employment, housing, credit, public education anti-discrimination laws and to Massachusetts' hate crimes law. All of these laws also protected several other characteristics, including sexual orientation, disability, sex, age, race, ancestry and religion. The law went into effect on July 1, 2012.

Chapter 199 defines "gender identity" as "a person's gender-related identity, appearance or behavior, whether or not that gender-related identity or behavior is different from that traditionally associated with the person's physiology or assigned sex at birth."¹⁰ The law allows a person to demonstrate his/her gender identity by providing evidence including: medical history; care or treatment of the gender identity; consistent and uniform assertion of the gender identity; or any other evidence that the gender identity is sincerely held as part of a person's core identity."¹¹

⁸ This was not the first time that the Division had examined whether insurance carriers were acting in a discriminatory manner with respect to coverage under their insurance policies. In 1988, the Commissioner of Insurance had issued regulations prohibiting a life insurer from considering gender-based mortality differences in the underwriting of life insurance, and which provided that "[n]o policy...shall, on the basis of...sex...treat any covered person...differently than it treats or would treat any other covered person...." 211 CODE MASS. REGS. § 35.04 (2) (1987). In *Telles v. Commissioner of Insurance*, the Supreme Judicial Court held that the Commissioner "lacked either express or implied authority to promulgate the regulations," and that "the regulations at issue directly conflict with several of the statutes which regulate insurance practices." The Court concluded that the regulations were "void because the commissioner lacked authority to issue the regulations. *See* 401 Mass. 560, 565-566 (1991).

⁹ MASS. GEN. LAWS ch. 199 (2011).

¹⁰ *Id.* at § 1.

¹¹ *Id.*

Chapter 199, while formally amending various laws precluding discrimination in employment, housing and other areas on the basis of one's "gender identity," specifically did not amend any laws covering discrimination in the areas of health insurance law. At that time, however, several other states had amended their respective insurance laws to specifically preclude discrimination in health insurance on account of a person's gender identity or because of a person's gender dysphoria.

For example, in California, the regulations governing health insurance companies had been specifically amended to require that an admitted health insurer could not "discriminate on the basis of an insured's or prospective insured's gender identity, or on the basis that the insured or prospective insured is a transgender person."¹²

The discrimination prohibited by California regulation includes "[d]enying, cancelling, limiting or refusing to issue or renew an insurance policy on the basis of an insured's or prospective insured's actual or perceived gender identity, or for the reason that the insured or prospective insured is a transgender person."¹³

In addition, the California regulation prohibits health carriers from:

[d]enying or limiting coverage, or denying a claim, for services...due to an insured's actual or perceived gender identity or for the reason that the insured is a transgender person [including]:
(1) Health care services related to gender transition if coverage is available for those services under the policy when the services are not related to gender transition, including but not limited to hormone therapy, hysterectomy, mastectomy, and vocal training; or (2) Any health care services that are ordinarily or exclusively available to individuals of one sex when the denial or limitation is due only to the fact that the insured is enrolled as belonging to the other sex or has undergone, or is in the process of undergoing, gender transition.¹⁴

¹² See CAL. CODE REGS. tit. 10, § 2561.2 (2012).

¹³ *Id.* at § 2561.2(a)(1).

¹⁴ *Id.* at § 2561.2(a)(4)(A) - (B).

In Oregon, the Insurance Division of the Department of Consumer and Business Services issued Bulletin INS 2012-01 in 2012.¹⁵ This bulletin stated that a health insurer in the state cannot discriminate in providing coverage on the basis of an insured's or prospective insured's gender identity or gender dysphoria.¹⁶ The Oregon Insurance Division stated that the bulletin was designed to provide guidance to health insurers about how to conform to provisions of the Oregon Equality Act of 2007, in which "sexual orientation" is defined to include an individual's actual or perceived gender identity, "regardless of whether the individual's gender identity, appearance, expression or behavior differs from that traditionally associated with the individual's sex at birth."¹⁷

The Oregon Insurance Division noted that because the Oregon insurance code already prohibited discrimination in the provision of health insurance coverage on the basis of "sexual orientation," health carriers could not deny or limit coverage or deny a claim for a procedure provided for gender identity or gender dysphoria if the same procedure were allowed in the treatment of another medical condition.¹⁸ Although a health insurer could categorically exclude coverage for a particular condition or treatment, the insurer could not base such exclusion on gender identity.¹⁹

In Vermont, the Department of Financial Regulation, Division of Insurance issued Insurance Bulletin No. 174 in 2013, which provides that notice to insurers that health care plans could not exclude coverage for medically necessary services for transgender people, including gender reassignment surgeries.²⁰ The bulletin rested specifically on the 2007 Vermont law, Act 41, which specifically prohibits discrimination on the basis of "gender identity."²¹ The bulletin noted that the law prohibiting gender identity discrimination applied to insurance companies, and as such, effective January 1, 2014, the Vermont Division of Insurance precluded

¹⁵ See OR. INS. DIVISION, Bulletin INS 2012-01 (Dec. 19, 2012) <https://dfr.oregon.gov/laws-rules/Documents/Bulletins/bulletin2012-01.pdf>. [hereinafter *Bulletin 2012-01*].

¹⁶ *Id.* at 1.

¹⁷ See *id.*, citing OR. REV. STAT. § 174.100.

¹⁸ *Id.* at 3.

¹⁹ *Id.* at 3-4.

²⁰ VT. D.F.R., DIVISION OF INS., Ins. Bulletin No. 174 (Apr. 22, 2013), <https://dfr.vermont.gov/sites/finreg/files/regbul/dfr-bulletin-insurance-174.pdf>. [hereinafter *Bulletin No. 174*]

²¹ See VT. STAT. ANN. tit. 1, § 144 (2007).

health insurers from excluding from coverage care related to gender transition.²²

Unlike in these other states, however, the 2011 law in Massachusetts, Chapter 199 only precluded discrimination in employment, housing and other areas on the basis of one's "gender identity;" it did not explicitly extend to precluding the exclusion of gender transition-related medical care from health insurance policies.²³ As such, the Division determined that it needed to look elsewhere to see if there was any other basis in Massachusetts law or court decisions for disallowing such exclusions.

During the time period when the Division was conducting its review, a new decision was handed down by the U.S. District Court for the District of Massachusetts concerning issues related to gender dysphoria. In *Kosilek v. Spencer*,²⁴ the District Court had held that a prisoner's gender identity disorder constituted a serious medical need that triggered Eighth Amendment protection.²⁵ In making its decision, the District Court was presented with testimony from Department of Correction ("DOC") physicians, who testified that "Kosilek is now suffering a degree of mental anguish that itself constitutes a serious harm that requires adequate treatment."²⁶

²² *Bulletin No. 174, supra* note 20.

²³ Chapter 199 amended various chapters of the Massachusetts General Laws, but none related to insurance. *See e.g.* MASS. GEN. LAWS ch. 199 § 1.

²⁴ 889 F. Supp.2d 190 (D. Mass. 2012) *aff'd*, 740 F.3d 733 (1st Cir. 2014), *reh'g en banc granted, opinion withdrawn* (Feb. 12, 2014), *on reh'g en banc*, 774 F.3d 63 (1st Cir. 2014), and *rev'd*, 774 F.3d 63 (1st Cir. 2014).

²⁵ *Id.* The decision was initially affirmed by the First Circuit, but on February 12, 2014, the First Circuit agreed to hear the case *en banc* and withdrew their initial opinion.

²⁶ *See id.* at 229. While the court in *Kosilek* used the term "gender identity disorder," the American Psychiatric Association changed the term "gender identity disorder" to "gender dysphoria" in the then latest version of the Diagnostic and Statistical Manual of Mental Disorders ("DSM")—DSM V—in December 2012, in order to "respect the individuals identified by offering a diagnostic name that is more appropriate to the symptoms and behaviors they experience without jeopardizing their access to effective treatment options." *See* Am. Psychiatric Ass'n, *Gender Dysphoria* (2013), <https://www.ca1.uscourts.gov/sites/ca1/files/citations/Gender%20Dysphoria%20Fact%20Sheet%20202.pdf>. The terms "gender dysphoria," "gender

The District Court in *Kosilek* ordered the DOC to provide the means for Kosilek to undergo gender reassignment surgery. In making its ruling the court relied on the fact that “[a]ll of the doctors who testified at trial, except for [one], provided evidence that sex reassignment surgery for Kosilek is both medically necessary and the only adequate treatment for his severe gender identity disorder.”²⁷ Without such surgery, the court found Kosilek was at a high risk of further attempts at suicide.²⁸

The *Kosilek* court, however, limited its holding to the prison context, and noted that the U.S. Constitution’s Eighth Amendment imposes certain duties on prison officials to provide humane conditions of confinement, adequate food, clothing, shelter, and medical care.²⁹ The District Court cited to the Supreme Court’s view on a state’s duties to prisoners under the Eighth Amendment: “[t]o incarcerate, society takes from prisoners the means to provide for their own needs. Prisoners are dependent on the state for food, clothing, and necessary medical care. A prison’s failure to provide sustenance for inmates may actually produce physical torture or a lingering death.”³⁰

Ultimately, however, the Division did not find that the *Kosilek* court’s determination, which was based on the court’s conclusion that the Department of Corrections had violated the Constitution’s prohibition of cruel and unusual punishment, was instructive in answering the question as

identity disorder,” and “transsexualism” were often used interchangeably by courts. *See e.g.*, *South v. Gomez*, No. CV-95-01070-DFL at *1 (9th Cir. Feb. 25 2000) (Westlaw) (noting that “gender dysphoria [is] more commonly known as transsexualism”); *see also* *Glenn v. Brumby*, 724 F. Supp. 2d 1284, 1290 n.5 (N.D. Ga. 2010) *aff’d* 663 F.3d 1312 (11th Cir. 2011) (“[Gender identity disorder (GID)] and transsexualism are closely related and are sometimes used as synonyms....”).

²⁷ *See Kosilek*, 889 F.Supp. 2d at 233.

²⁸ *Id.*

²⁹ *Id.* at 203. The District Court noted that “a prison official acts with deliberate indifference and violates the Eighth Amendment if, knowing of a real risk of serious harm, she denies adequate treatment for a serious medical need for a reason that is not rooted in the duties to manage a prison safely and to provide the basic necessities of life in a civilized society for the prisoners in her custody.”

³⁰ *See id.* at 198, *citing* *Brown v. Plata*, 131 S.Ct. 1910, 1928 (2011).

to whether a private insurance carrier would violate Massachusetts law when the carrier excluded coverage for gender transition-related treatment.³¹

Therefore, the Division began to explore whether there was any other basis in federal and state law for prohibiting health insurance carriers from excluding from coverage medical treatment for persons with gender dysphoria.

II. MENTAL HEALTH PARITY

One area that the Division examined was whether the exclusion of gender transition-related medical care from health insurance policies in Massachusetts might amount to unlawful discrimination based on a person's mental health under the Massachusetts mental health parity law.³² The Massachusetts mental health parity law required that insurance plans cover mental health benefits on a non-discriminatory basis for the medically necessary treatment of any "mental disorder" listed in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders ("DSM").³³ The Massachusetts mental health parity law provides that: "[a]n individual policy of accident and sickness insurance... shall provide mental health benefits on a nondiscriminatory basis to residents of the commonwealth...for the diagnosis and medically necessary and active treatment of any mental disorder, as described in the most recent edition of the DSM, that is approved by the commissioner of mental health."³⁴

The Division looked to the state of Connecticut, where the Connecticut Division of Insurance in 2013 in its Bulletin IC-34, relied upon the state's mental health parity statute as the basis for concluding that the exclusion of gender transition-related medical care from health insurance

³¹ *See id.* at 205. The Kosilek court's finding that a prisoner completely relied on the state for medical care was a key rationale supporting the Court's decision that by *not* treating a prisoner for her gender dysphoria, the state had violated the Constitution. As the District Court noted, it "has long been well-established that it is cruel for prison officials to permit an inmate to suffer unnecessarily from a serious medical need. It is unusual to treat a prisoner suffering severely from a gender identity disorder differently than the numerous inmates suffering from more familiar forms of mental illness."

³² MASS. GEN. LAWS ch. 175 § 47B (a) (2015).

³³ *Id.*

³⁴ *See* MASS. GEN. LAWS ANN. ch. 175, § 47B (a).

policies in Connecticut was impermissible.³⁵ The Connecticut mental health parity statute provides that “[e]ach individual health insurance policy...shall provide benefits for the diagnosis and treatment of mental or nervous conditions.”³⁶ The Connecticut bulletin further stated that the Connecticut mental health parity statute,³⁷ in conjunction with the Connecticut group health insurance statute, together “require health insurers to pay ‘covered expenses’ for treatment provided to individuals with gender dysphoria where treatment is deemed necessary under generally accepted medical standards.”³⁸

The language in the Connecticut mental health parity statute mirrors that in the Massachusetts statute, which prohibits an insurer from “provid[ing] mental health benefits on a discriminatory basis to residents of the commonwealth...for the diagnosis and medically necessary and active treatment of any mental disorder, as described in the most recent edition of the DSM, that is approved by the commissioner of mental health.”³⁹ Therefore, at the time, the Division considered whether perhaps under Massachusetts mental health parity law, as in Connecticut, an argument could be made that if an individual is diagnosed with gender dysphoria, as recognized in the latest DSM as a “mental disorder,” an insurer could be prohibited from limiting or withholding coverage for medically necessary treatment, where the insurer would provide the same treatment to individuals who require it for a different medically necessary reason.

The Division ultimately concluded that it would not rely on the Massachusetts mental health parity laws as the basis for concluding that the exclusion of gender transition-related medical care from health insurance policies in the state was not permissible because of the continued debate within the activist community as to whether being a transgender person was a “mental disorder” at all.

As noted above, in December 2012, the American Psychiatric Association announced that it approved changes in its official manual for classifying mental illnesses, known as DSM-5, formally eliminating the term “gender identity disorder,” and replacing it with the term “gender

³⁵ See CONN. DIV. INS., Bulletin IC-34 (Dec. 19, 2013), <https://portal.ct.gov/-/media/CID/BulletinIC37GenderIdentityNondiscriminationRequirementspdf.pdf?la=en>.

³⁶ See CONN. GEN. STAT. § 38a-488a (b) (2013).

³⁷ *Id.*

³⁸ See Bulletin IC-34 at 1.

³⁹ See MASS. GEN. LAWS ch. 175, § 47B(a).

dysphoria.”⁴⁰ The term “gender identity disorder” had been long considered stigmatizing by mental health specialists and lesbian, gay, bisexual and transgender activists.⁴¹ “Gender dysphoria” instead focuses the attention on only those who feel distressed by their gender identity.⁴² At the time of the change in terms in the DSM-5, there had been calls by activists to remove the diagnosis altogether just as homosexuality had been removed from the DSM in 1973, but gender dysphoria was ultimately left as a diagnosis to ensure that a transgender person could still access health care if needed.⁴³

While many transgender activists felt that the gender dysphoria diagnosis remains a “powerful legal tool” when challenging discrimination in health insurance plans and services, other activists disagreed, stating that the new DSM criteria did not go nearly far enough in clarifying that “nonconformity to birth-assigned roles and victimization from societal prejudice do not constitute mental pathology,” and that being a transgender person was not a mental disorder.⁴⁴ The advocacy organization GLAAD noted similar concerns at the time, stating that:

Some transgender advocates see this approved change in the DSM-V as an important step toward removing stigma against transgender people based on false stereotypes about gender identity and expression, as well as the word “disorder.” Transgender people may no longer be subject to a lifelong default diagnosis of their mental health.... However, other transgender advocates note the barriers this change may create to accessing trans-related medical care, which could already be difficult to access and prohibitively expensive even before the change.⁴⁵

⁴⁰ *Supra* note 26.

⁴¹ See Moni Basu, *Being Transgender No Longer a Mental 'Disorder' in Diagnostic Manual*, CNN (Dec. 27, 2012, 10:46 AM), <http://inamerica.blogs.cnn.com/2012/12/27/being-transgender-no-longer-a-mental-disorder-in-diagnostic-manual/> (“CNN Report”).

⁴² *Id.*

⁴³ *Id.*

⁴⁴ *Id.*

⁴⁵ See Dani Heffernan, *The APA Removes “Gender Identity Disorder” from Updated Mental Health Guide*, GLAAD (Dec. 3, 2012), <https://www.glaad.org/blog/apa-removes-gender-identity-disorder-updated-mental-health>

Therefore, while the Connecticut Insurance Department relied upon its mental health parity law to establish the principle that excluding coverage treating gender dysphoria would be a parity violation, this conclusion necessitated a finding that gender dysphoria was a major mental disorder subject to a mental health parity analysis. The Division, however, did not believe that it was appropriate to reach a similar conclusion, because there was no strong consensus in favor of this position in the transgender community in Massachusetts, and there were many transgender persons who strongly believed that being transgender was not a mental disorder or pathology. As such, the Division concluded that it could not rely upon Massachusetts mental health parity law to preclude carriers from excluding coverage for treating gender dysphoria.

III. UNFAIR INSURANCE PRACTICE

The Division next looked to whether excluding coverage for gender transition-related medical treatment from people with gender dysphoria violated Massachusetts unfair insurance practices law. In Massachusetts, unfair insurance practices governed under Massachusetts General Law Chapter (“Chapter”) 176D are considered the “making or permitting any unfair discrimination between individuals of the same class and of essentially the same hazard in the amount of premium, policy fees, or rates charged for any policy or contract of accident or health insurance or in the benefits payable thereunder, or in any of the terms or conditions of such contract, or in any other manner whatever.”⁴⁶

Thus, the argument for applying this law to the coverage issue at hand was that Chapter 176D, § 3(7)(b) would be applicable to individuals who require treatment for gender dysphoria because they are of the same class and of essentially the same hazard as individuals who require the same treatment for a different medically necessary reason. The first issue that was looked at was whether the two groups were of the “same class.” The Massachusetts statute, however, does not define “class.”

In *Life Insurance Association of Massachusetts v. Commissioner of Insurance*,⁴⁷ the Massachusetts Supreme Judicial Court struck down a Division of Insurance regulation which prohibited underwriting practices of insurers regarding the testing of prospective insureds for exposure to HIV. The Court noted that the “basic principle underlying statutes [like Chapter

h-Guide.

⁴⁶ See MASS. GEN. LAWS ANN. Ch. 176D, § 3(7)(b) (West 2012).

⁴⁷ 530 N.E.2d 168 (1988).

176D § 3]...is that insurers have the right to classify risks and to elect not to insure risks if the discrimination is fair.”⁴⁸

The Court also noted that the intended result of the process is that persons of substantially the same risk will be grouped together, paying the same premiums, and will not be subsidizing insureds who present a significantly greater hazard. The Court found that insurers, under Chapter 176D § 3, have a general right to discriminate fairly. The Court also noted: “[i]t is not seriously denied that persons who have HIV antibodies, as a group, are at greater risk of illness and have shorter life expectancies than those who do not have HIV antibodies.”⁴⁹ The Court’s ruling indicates that it did not consider persons who present greater risks to the insurer (individuals with HIV) to be in the “same class” as those who present lesser risks (individuals without HIV).

The ruling in *Life Ins. Ass’n of Massachusetts* was reinforced by the SJC in *Telles v. Commissioner of Insurance*.⁵⁰ The question in *Telles* was whether the Commissioner of Insurance could “lawfully issue regulations which prohibit life insurers from considering gender-based mortality differences in the underwriting of life insurance.”⁵¹ The Court noted that the Commissioner’s “unisex” regulation required individuals from different risk classes—males and females—to be grouped together.

Relying on *Life Ins. Ass’n of Massachusetts*, the Supreme Judicial Court in *Telles* found that requiring insurers to group men and women together, individuals typically in different risk classes, to be “in direct conflict” with Chapter 176D §3(7). In *Telles*, the Court held that the Commissioner of Insurance was without authority to promulgate regulations prohibiting life insurers from considering gender-based mortality differences in the underwriting of life insurance, insurers had the statutory right to classify risks. Thus, gender-based classifications for the determination of insurance rates were permitted under the statutory scheme.

The *Telles* court read the “same class” language to mean that “insureds must be treated in accordance with their risk classification.”⁵² As such, the *Telles* court would likely interpret the “same class” language to mean “same risk classification,” and if two groups present different risks to

⁴⁸ See *id.* at 171.

⁴⁹ See *id.*

⁵⁰ 574 N.E.2d 359 (1991).

⁵¹ *Id.* at 360.

⁵² 574 N.E.2d at 361.

the insured, the groups would be considered to be in different classes for purposes of Chapter 176D § 3.

At the time, the Division noted that the holdings in *Life Ins. Ass'n of Massachusetts* and *Telles* might be distinguishable from the question of whether an insurer can exclude coverage for medically necessary treatment from individuals solely because they have gender dysphoria. *Life Ins. Ass'n of Massachusetts* and *Telles* dealt with individuals who were in different risk classifications: individuals with and without HIV; and men and women. An individual with gender dysphoria and an individual with cervical cancer may both require a hysterectomy as part of their medically necessary treatment, and as such, could be viewed as being in the same risk classification.

The costs and risks these two groups present to the insurer would be the same—the cost of the hysterectomy, for example—even though the needs for the treatments have different causes. Since an individual with gender dysphoria would not necessarily be costlier than an individual who requires the same treatment for a different medically necessary reason, these two groups would likely be placed in the “same class,” and *Life Ins. Ass'n of Massachusetts* and *Telles* decisions would not necessarily prevent the Division from prohibiting discrimination between the two groups. Therefore, if an insurer denies coverage for a particular treatment only to individuals with gender dysphoria, but not to individuals who need the same medical treatment for a different reason, then the insurer might be in violation of Chapter 176D § 3.

To interpret the term “same class” to include individuals with and without gender dysphoria would have aligned the Division with the approach taken by the state of Colorado. Colorado’s Division of Insurance treated individuals with and without gender dysphoria as belonging to the same class for purposes of the Colorado unfair insurance practices statute. The Colorado Division of Insurance issued a bulletin prohibiting discrimination against individuals with gender dysphoria based, in part, on their counterpart to the Massachusetts unfair insurance practices law.⁵³

⁵³ See COLO. DIV. INS., BULL. NO. B-4.49, INSURANCE UNFAIR PRACTICES ACT PROHIBITIONS ON DISCRIMINATION BASED UPON SEXUAL ORIENTATION (2013), <http://www.one-colorado.org/wp-content/uploads/2013/03/B-4.49.pdf>. As noted in the bulletin, Colorado law defined “sexual orientation” as “a person’s orientation toward heterosexuality, homosexuality, bisexuality, or transgender status or another person’s perception thereof” and such definition applied to every statute, including the unfair insurance practices law.

The Colorado unfair insurance practices law prohibits any unfair discrimination “between individuals of the same class...in the amount of premium, policy fees, or rates charged for any policy of sickness and accident insurance, in the benefits payable under such policy, in the terms or conditions of the policy, or in any other manner.”⁵⁴ Although nothing in the Colorado Bulletin expressly states so, it appears likely that Colorado would consider individuals—with and without gender dysphoria—who require the same medically necessary treatment to be individuals of the “same class and of essentially the same hazard.”⁵⁵

Similarly, the D.C.’s Department of Insurance, Securities and Banking issued two bulletins in 2013 and 2014 respectively that prohibited gender identity discrimination. These bulletins were based on the District’s Unfair Insurance Trade Practices Act, which prohibited discrimination in health insurance based on gender identity or expression.⁵⁶ In its bulletin issued on February 27, 2014 (“February 2014 Bulletin”) prohibiting discrimination against individuals with gender dysphoria, the Department of Insurance, Securities and Banking articulated the interpretation of “same class and of essentially the same hazard” language the same way as Colorado’s Division of Insurance.

⁵⁴ See COLO. REV. STAT. 10-3-1104(1)(f)(XIII) (2018).

⁵⁵ In the case *Cortez v. Progressive County Mut. Ins. Co.*, No. 03-99-00846-CV (Sept. 13, 2001), the Texas Court of Appeals was looking at identical language contained in the Texas unfair insurance practices law, and concluded that the interpretation of “same class and of essentially the same hazard” language meant looking at the “treatment of the plaintiffs in comparison to other similarly situated individuals.” As such, it would be reasonable in Colorado to view individuals who require the same medically necessary treatment to be “similarly situated individuals.”

⁵⁶ See D.C. DEP’T OF INS., SEC. AND BANKING, BULL. NO. 13-IB-01-30/15, PROHIBITION OF DISCRIMINATION IN HEALTH INSURANCE BASED ON GENDER IDENTITY OR EXPRESSION (2013), <https://disb.dc.gov/publication/13-ib-01-3015-prohibition-discrimination-health-insurance-based-gender-identity-or>; see also D.C. DEP’T OF INS., SEC. AND BANKING, BULL. NO. 13-IB-01-30/15 Revised, PROHIBITION OF DISCRIMINATION IN HEALTH INSURANCE BASED ON GENDER IDENTITY OR EXPRESSION (2014), <https://disb.dc.gov/publication/disb-bulletin-13-ib-01-3015-revised-prohibition-discrimination-health-insurance-based>.

The D.C.'s February 2014 Bulletin cites the D.C.'s counterpart to the Massachusetts unfair competition in insurance statute, the District's Code § 31-2231.11.⁵⁷ The February 2014 Bulletin first clarifies that gender dysphoria is "a recognized medical condition under health insurance policies covering medical and hospital expenses, regardless of whether explicitly referenced."⁵⁸ Next, the February 2014 Bulletin noted the unfair competition statute applied to health insurance.⁵⁹

The District of Columbia's unfair competition statute varies slightly from that of Massachusetts' in that the statute expressly prohibits discrimination based on gender identity or expression. The District of Columbia's February 2014 bulletin went on to state that "[t]he only interpretive question that remains... is whether gender dysphoria diagnosed individuals and non-gender dysphoria diagnosed individuals seeking health insurance are 'of the same class and essentially the same hazard.'"⁶⁰

Because both sets of individuals were seeking coverage under the same health insurance policies offering benefits and services for recognized medical conditions, the District of Columbia's Department of Insurance, Securities and Banking in the bulletin concluded that for purposes of § 31-2231.11(b), the individuals were of the "same class" and "essentially the same hazard."⁶¹ To come to the conclusion reached by the District of Columbia, it does not appear necessary to have express language prohibiting discrimination based on gender identity or expression contained within the unfair insurance practices law, but the express language served to bolster the analysis. By concluding that individuals with and without gender dysphoria are of "the same class and essentially the same hazard" the District of Columbia appeared to agree with the state of Colorado.

Thus, the key issue for the Division in 2014 was whether it was reasonable to conclude that Massachusetts, like Colorado and the District of Columbia, would consider individuals with and without gender dysphoria who require medically necessary treatment to be individuals of the "same class and of essentially the same hazard." Only if the two groups were treated as being in the same class and essentially the same hazard, would Chapter 176D § 3 prohibit an insurer from "any unfair discrimination...in any...manner whatever" against individuals with gender dysphoria.

⁵⁷ See D.C. CODE § 31-2231.11 (2012).

⁵⁸ See Bulletin 13-IB-01-30/15, Revised at 2.

⁵⁹ *Id.* at 3.

⁶⁰ *Id.*

⁶¹ *Id.*

IV. SEXUAL ORIENTATION

In early 2014, the Division also looked at whether health insurers that excluded coverage for people with gender dysphoria violated Massachusetts laws prohibiting discrimination on the basis of sexual orientation. At the time, Massachusetts law generally prohibited sexual orientation discrimination in the areas of employment, housing, public accommodations, credit and services, and education as well as insurance.⁶² In the employment context, Massachusetts law unambiguously defined “sexual orientation” as including only “heterosexuality, bisexuality, or homosexuality.”⁶³ There was no specific Massachusetts statute or regulation, that specifically defined sexual orientation as “gender identity” or gender dysphoria.

Despite the lack of express statutory or regulatory authority to including individuals with gender dysphoria in the “sexual orientation” group, at least one Massachusetts court had issued an opinion that supported a broad interpretation of the meaning of “sexual orientation” discrimination. In 2002, in *Lie v. Sky Publishing Corporation*,⁶⁴ the Massachusetts Superior Court found that those who transgress traditional gender roles and defy stereotypes associated with their biological sex are less likely to be perceived as heterosexual than the general population.⁶⁵

As a result, the court held that the conflation of one’s appearance with one’s sexual orientation might lead to discrimination actionable under Chapter 151B’s definition of sexual orientation discrimination.⁶⁶ It did not appear at the time, however, that this interpretation was generally accepted in Massachusetts. Moreover, the court’s conclusion was at odds with the long-recognized differences between sexual orientation and gender identity,

⁶² See MASS. GEN. LAWS ch. 151B, § 1 (2004).

⁶³ See ch. 151B, § 3(6) (2012).

⁶⁴ No. 013117J (Mass. Super. Ct, Oct. 7, 2002).

⁶⁵ *Id.*

⁶⁶ See *Sky Publishing Corp.*, slip op. at 4, (citing *Rosa v. Park West Bank & Trust*, 214 F.3d 213, 216 (1st Cir. 2000)) (“It is...reasonable to infer...that [the teller] refused to give [the plaintiff] the loan application because she thought he was gay, confusing sexual orientation with cross-dressing”).

which lead to the conclusion that sexual orientation protections would not apply per se to protect individuals who were transgender.⁶⁷

V. SEX DISCRIMINATION

Another argument that the Division considered in early 2014 to preclude health insurers from excluding individuals from coverage for certain medical treatments because they have gender dysphoria, was that such an exclusion violates federal and Massachusetts laws which prohibit discrimination based on sex. In the absence of statutory language that defined the term “sexual discrimination” in health insurance laws as specifically including discrimination based upon “gender identity,” whether the term “sex discrimination” extended to protect individuals with gender dysphoria depended on the scope given to the term. Under a broad interpretation of the term, “sex discrimination” could include discrimination based on gender non-conformance and applies to individuals with gender dysphoria, while a narrow interpretation of sex discrimination would limit the term to include only discrimination based on an individual’s biological sex.

In 1989, the Supreme Court of the United States interpreted the term “sex discrimination” broadly in suits brought under Title VII of the Civil Rights Act. In *Price Waterhouse v. Hopkins*,⁶⁸ a plurality of the Court addressed sex discrimination in a suit brought by a female partnership candidate in an accounting firm who alleged she was discriminated against for appearing too “macho.”⁶⁹ In its ruling, the Court moved away from the traditional, limited view of sex discrimination and stated: “we are beyond the day when an employer could evaluate employees by assuming or insisting that they matched the stereotype associated with their group, for ‘[i]n forbidding employers to discriminate against individuals because of their sex, Congress intended to strike at the entire spectrum of disparate treatment

⁶⁷ See *Sexual Orientation and Gender Identity Definitions*, HUMAN RIGHTS CAMPAIGN, <https://www.hrc.org/resources/sexual-orientation-and-gender-identity-terminology-and-definitions> (defining “gender identity” as the “innermost concept of self as male, female, a blend of both or neither – how individuals perceive themselves and what they call themselves,” while “sexual orientation” is the “inherent or immutable enduring emotional, romantic or sexual attraction to other people”).

⁶⁸ See 490 U.S. 228 (1989).

⁶⁹ See *id.* at 235.

of men and women resulting from sex stereotypes.”⁷⁰ The Court concluded that the term “sex discrimination” could include discrimination against persons who fail to conform to gender stereotypes.⁷¹

In *Smith v. City of Salem Ohio*,⁷² the Sixth Circuit relied on *Price Waterhouse* to expressly recognize a cause of action for a transgender person claiming protection under Title VII. The *Smith* case involved a city fire department employee, who was born biologically male and was diagnosed with gender identity disorder while working for the city fire department. After the city pressured the employee to submit to multiple psychiatric evaluations by doctors of their choosing, the employee brought a Title VII action alleging sex discrimination.⁷³

The Sixth Circuit noted that pre-*Price Waterhouse* federal courts routinely rejected expanding the definition of “sex” to include gender non-conforming individuals, but that those cases had been “overruled by the logic and language” of *Price Waterhouse*. The court ultimately held that allegations of discrimination based upon the employee’s gender non-conforming behavior and appearances were actionable pursuant to Title VII.⁷⁴ Post *Price Waterhouse* and *Smith*, under Title VII, the term “sex” appeared to encompass both biological sex and the failure to conform to socially prescribed gender expectations.

The First Circuit had similarly interpreted “sex discrimination” as being broad in scope. This is evident in the *Rosa v. Park West Bank & Trust Co.*,⁷⁵ decision. In that case, the court found that discrimination based on an individual’s habit of cross dressing may be considered sex discrimination. In *Rosa*, the First Circuit concluded that a biological male who presented and lived as a female may be able to establish a cause of action for sex discrimination under the Equal Credit Opportunity Act (“ECOA”), which prohibits discrimination with respect to any aspect of a credit transaction on the basis of sex, where she was denied a loan application from a bank because of her feminine attire.⁷⁶

⁷⁰ *See id.* at 251 (quoting *Sprogis v. United Air Lines, Inc.*, 444 F.2d 1194, 1198 (7th Cir. 1971)).

⁷¹ 490 U.S. at 251.

⁷² 378 F.3d 566 (6th Cir. 2004).

⁷³ *Id.* at 568–70.

⁷⁴ *Id.* at 573–75, 578.

⁷⁵ 214 F.3d 213 (1st Cir. 2000).

⁷⁶ *Id.* at 215–16.

The court found it reasonable to infer that the Bank told “Rosa to go home and change because [the bank] thought that Rosa’s attire did not accord with his male gender: in other words, that Rosa did not receive the loan application because he was a man, whereas a similarly situated woman would have received the loan application.”⁷⁷ The court cited the Supreme Court in *Price Waterhouse* to support the conclusion that “stereotyped remarks [including statements about dressing more ‘femininely’] can certainly be evidence that gender played a part” in the discrimination.⁷⁸

The broad interpretation of “sex discrimination” had also been extended to cases where the discrimination was based on an employee’s perceived homosexuality. In *Centola v. Potter*,⁷⁹ the U.S. District Court in Massachusetts held an employee’s Title VII sex discrimination claim could survive a motion for summary judgment where the employee was subject to “constant” harassment which focused on his being homosexual.⁸⁰ The district court found that the employee’s “[c]o-workers and supervisors discriminated against him because he failed to meet their gender stereotypes of what a man should look like, or act like.

In so doing, they created an objectively hostile and abusive work environment in violation of Title VII.”⁸¹ The district court relied on *Price Waterhouse* and *Rosa* when it held that: “If an employer acts upon stereotypes about sexual roles in making employment decisions, or allows the use of these stereotypes in the creation of a hostile or abusive work environment, then the employer opens itself up to liability under Title VII’s prohibition of discrimination on the basis of sex.”⁸²

The broad interpretation of sex had been accepted at the time by the Massachusetts Superior Court in *Doe ex rel. Doe v. Yunits*,⁸³ which addressed whether a school policy which prevented a male student from dressing in attire typically associated with females was illegal sex discrimination. Also citing *Price Waterhouse*, the court held that the school’s policy constituted sex discrimination under Chapter 76, § 5 (the school attendance discrimination statute), because the school prevented the student

⁷⁷ *Id.* at 215.

⁷⁸ *Id.* at 216 (citations omitted).

⁷⁹ *Centola v. Potter*, 183 F. Supp. 2d 403 (D. Mass. 2002).

⁸⁰ 183 F. Supp. 2d at 407.

⁸¹ *Id.* at 409.

⁸² *Id.*

⁸³ *See Doe ex rel. Doe v. Yunits*, 2000 WL 33162199 (Mass. Super. Oct. 11, 2000).

from attending school in clothing associated with the female gender solely because the student was male.⁸⁴

Similarly, at the administrative level, the Massachusetts Commission Against Discrimination (“MCAD”) found that discrimination based on an individual’s transgender status was actionable under the Massachusetts unfair discrimination in employment statute, Chapter 151B, as “sex discrimination.”⁸⁵ In proceedings before the MCAD, the employee, a transgender woman, alleged that her supervisor discriminated against her because of her sex. The employee alleged the supervisor had issued pretextual written warnings for insubordination and threatened her with termination after the employee complained about her supervisor’s harassing behavior towards her.⁸⁶

The MCAD, also citing *Price Waterhouse*, found that “[s]ex discrimination is a concept that is read broadly; in other words, illegal “sex discrimination” takes into account non-anatomical concepts, like gender.”⁸⁷ The MCAD ultimately held that “sex discrimination, as prohibited by chapter 151B, includes a prohibition against discrimination against transgender individuals.”⁸⁸

When applying the above analysis to the insurance context, the Division looked at whether excluding individuals with gender dysphoria from coverage for certain medical treatments would constitute discrimination based on stereotyped notions of appropriate gender behavior. Based on the reasoning in the above-referenced authority, the Division concluded that Massachusetts courts would follow the majority of courts that had found that a broad interpretation should be given to “sex discrimination.”

Therefore, the Division determined that if an insurer refused to cover gender assignment-related medical treatment because the insured failed to conform to the insurer’s idea of how a man or woman should look and behave, then the insured would have been discriminated against based on their sex. Thus, if a health insurer denied to provide coverage for medically

⁸⁴ *Id.* at *7

⁸⁵ *See* Millet v. Lutco, MCAD Docket No. 98 BEM 3695, 23 M.D.L.R. 231 (Oct. 10, 2001), <https://www.mass.gov/files/documents/2017/10/25/milletvlutco-2001.pdf>.

⁸⁶ *See id.* at 1.

⁸⁷ *See id.* at 2.

⁸⁸ *See id.* at 5.

necessary treatment based on an individual's gender dysphoria then this would be considered prohibited sex discrimination under Massachusetts law.

On the other hand, it's possible that people with gender dysphoria would be excluded from coverage not because the insurer had antiquated notions of what is appropriate behavior, but because the insurer believed the medical treatment being sought by the insured was experimental. Hypothetically, an insurer could exclude experimental surgeries from coverage, to a male or female, and not base the exclusion on the individual's sex. Such an explanation may constitute a valid reason for denying treatment. However, in this scenario the issue would be between the parties to address the legitimacy of the treatment, and not, as it is here, on whether a blanket exclusion of coverage relating to gender transition health care—where the same treatment is covered for other medically necessary reasons—would be a form of prohibited sex discrimination.

Further support for the broad interpretation of sex discrimination was found at the time with two federal agencies which had addressed the issue. In 2012, the U.S. Department of Health and Human Services' Office for Civil Rights ("OCR") had determined that "sex discrimination" is extended to claims of discrimination based on gender identity or failure to conform to stereotypical notions of masculinity or femininity.⁸⁹ Likewise, in 2012, the U.S. Equal Employment Opportunity Commission had issued a formal ruling that gender identity discrimination is per se "sex discrimination."⁹⁰ In addressing the scope of sex discrimination, these two federal agencies both adopted a broad interpretation of "sex discrimination" and extended it to provide protection from discrimination to those individuals with gender dysphoria.

The letter from Leon Rodriguez, Director of the OCR, to Maya Rupert, Federal Policy Director, National Center for Lesbian Rights, dated July 12, 2012 (the "Rodriguez Letter"), stated that under Federal law, gender identity was viewed as a protected class with respect to health care plans

⁸⁹ See *Letter from Leon Rodriguez, Director of Office of Civil Rights, Department of Health and Human Services ("HHS"), to Maya Rupert, Federal Policy Director, National Center for Lesbian Rights, dated July 12, 2012*, <http://www.washingtonblade.com/content/files/2012/08/101981113-Response-on-LGBT-People-in-Sec-1557-in-the-Affordable-Care-Act-from-the-U-S-Dept-of-Health-and-Human-Services.pdf> [hereinafter *Rodriguez Letter*].

⁹⁰ See *Macy v. Eric Holder, Atty General, U.S. Dept. of Justice, EEOC Appeal No. 0120120821 (April 24, 2012)*, http://www.law.columbia.edu/sites/default/files/microsites/gender-sexuality/macy-v-holder_edited.pdf.

under the Affordable Care Act (“ACA”). For example, Director Rodriguez noted that Section 1557 of the ACA specifically prohibited discrimination in health care programs on the basis of gender identity, race, color, national origin, sex, sex stereotypes, age or disability.⁹¹ As such, health insurers, hospitals, the health insurance exchanges, and any other entities that received federal funds are covered by this law.⁹²

As noted in the Rodriguez Letter, discrimination against transgender people in federal health programs or health programs that receive federal funds is prohibited under the ACA. The letter also notes that the Obama Administration had interpreted existing non-discrimination law — including Title VII of the Civil Rights Act of 1964 and Title IX of the Education Amendments of 1972 — to mean that the sex-discrimination protections under the ACA also applied to transgender people:

We agree that Section 1577’s sex discrimination prohibition extends to claims of discrimination based on gender identity or failure to conform to stereotypical notions of masculinity or femininity and will accept such complaints for investigation.... Section 1557 also prohibits sexual harassment and discrimination regardless of actual or perceived sexual orientation or gender identity of the individuals involved.⁹³

In the U.S. Equal Employment Opportunity Commission (“EEOC”) ruling in 2012 in *Macy v. Eric Holder*, the complainant, a transgender police detective in Phoenix, Arizona, had alleged employment discrimination in violation of Title VII of the Civil Rights Act of 1964. The EEOC found that gender identity and transgender discrimination was per se “sex discrimination” under Title VII.⁹⁴ The agency found that:

Title VII’s prohibition on sex discrimination proscribing gender discrimination, and not just discrimination on the basis of biological sex, is important. If Title VII proscribed only discrimination on the basis of biological sex, the only prohibited gender-based disparate treatment would be when an employer prefers a man over a woman,

⁹¹ *Rodriguez Letter* at 1.

⁹² *Id.*

⁹³ *Id.*

⁹⁴ *See Macy*, EEOC Appeal No. 0120120821 at 14.

or vice versa. But the statute's protections sweep far broader than that, in part because the term "gender" encompasses not only a person's biological sex but also the cultural and social aspects associated with masculinity and femininity.⁹⁵

The EEOC concluded its opinion by stating that "intentional discrimination against a transgender individual because that person is transgender is, by definition, discrimination "based on...sex" and such discrimination therefore violates Title VII."⁹⁶

If the Division were to find persuasive the reasoning in *Price Waterhouse, Smith, Rosa, Centola, Yunits, Macy* and the Rodriguez Letter, then there were several Massachusetts statutes which prohibited sex discrimination in the business of insurance that might be found broad enough to encompass discrimination in health insurance coverage against persons with gender dysphoria. For example, Chapter 175, § 24A provides:

No company authorized to issue policies of accident or sickness insurance, policies providing coverage against disability from injury or disease, or policies of life or endowment insurance shall refuse to issue such a policy or limit the coverages normally contained therein with respect to the risk of such loss solely because of the sex of the insured.⁹⁷

Therefore, excluding health insurance coverage for gender dysphoria-related treatment could be considered prohibited sex discrimination under existing Massachusetts law because it would be a limitation on coverage based "solely because of the sex of the insured."⁹⁸

VI. TRANSITIONING TO A NEW VIEW

As discussed above, at the beginning of 2014, the Division began to review Massachusetts' laws, as well as federal law and the law of other states, to determine whether health insurance carriers should be prohibited from excluding from coverage appropriate medical treatment for persons with gender dysphoria. We learned that while there was no Massachusetts statute or regulation that specifically prohibited health insurance carriers

⁹⁵ *Id.* at 6-7.

⁹⁶ *Id.* at 14

⁹⁷ MASS. GEN. LAWS ch. 175, § 24A (2018).

⁹⁸ *Id.*

from formally excluding coverage for persons with gender dysphoria for gender transition-related medical care including gender assignment surgery, hormone replacement therapy and other treatments, the Division did conclude that there were at least two possible bases for proscribing health insurers from excluding such coverage under their health plans.

One such possible basis was that excluding coverage for medically-necessary treatment for gender dysphoria would violate the Massachusetts unfair insurance practices law Chapter 176D. Making such a finding, however, would have required coming to the conclusion that individuals with gender dysphoria are of the “same class and of essentially the same hazard” as those without gender dysphoria. The Colorado Division of Insurance and the District of Columbia’s Department of Insurance, Securities and Banking had both come to this conclusion based upon their own unfair insurance practices laws.

But in the case of Colorado, Colorado law defined “sexual orientation” as “a person’s orientation toward heterosexuality, homosexuality, bisexuality, or transgender status or another person’s perception thereof” and this definition applied to the state’s unfair insurance practices law. With respect of the District of Columbia, the district’s unfair competition statute was different from that of Massachusetts’ statute in that the district’s statute expressly prohibited discrimination based on gender identity or expression, something that the Massachusetts unfair insurance practice statute did not do. As a result, the Division concluded that there wasn’t nearly as strong a case to be made in Massachusetts as in Colorado or the District of Columbia, because of the lack of any statutory law directly applying any protection for gender identity to the state’s unfair insurance practices law.

Nevertheless, the Division did determine that there was a very strong argument to be made for precluding health insurers from excluding individuals with gender dysphoria from coverage for certain medically necessary treatments would be the state’s existing prohibition on “sex discrimination” in the provision of health insurance, based on stereotyped notions of appropriate gender behavior. In this regard, the Division would be following the lead of the majority of courts which had concluded that a broad interpretation should be given to the term “sex discrimination.”

Under this analysis, if an insurer refused to cover medically necessary treatment because the insured failed to conform to the insurer’s idea of how a man or woman should look and behave, then the insured has been discriminated against based on his or her “sex.” Thus, the Division

concluded that denying medically necessary treatment based on an individual's gender dysphoria, and formally excluding from coverage for persons with gender dysphoria, gender transition-related medical care including gender assignment surgery, hormone replacement therapy and other treatments, must be considered prohibited sex discrimination under Massachusetts law.

In early June 2014, the Division came to the final conclusion that the denial of coverage by health insurance companies for gender transition-related medical care including gender assignment surgery, hormone replacement therapy and other treatments based on an individual's gender identity or gender dysphoria must be declared to be sex discrimination that was prohibited under Massachusetts law. As a result, Division issued its Bulletin 2014-03 on June 20, 2014.⁹⁹

VII. THE AFTERMATH

As a result of the issuance of the Bulletin, the previous nearly uniform exclusion from coverage of gender identity or gender dysphoria-related treatment by Massachusetts health plans became no longer permissible in the Commonwealth, as the Division determined that exclusions from coverage for gender transition-related medical care would no longer be allowed.¹⁰⁰ Once the Bulletin was issued, the health plans in the state immediately complied with its directives, and began to work with advocacy groups and state agencies to ensure that not only would coverage be available for gender transition-related medical treatment, but also that guidelines were developed to determine medical necessity for gender reassignment surgery.¹⁰¹

Since the issuance of the Division's issuance of the Bulletin in Massachusetts in 2014, the insurance departments of several other states issued insurance bulletins or guidance on the application of anti-discrimination laws to health insurance coverage for the treatment of gender dysphoria.¹⁰² The federal government was also moving ahead on the issue of protecting the rights of persons with gender dysphoria under federal law.

⁹⁹ See Bulletin 2014-03.

¹⁰⁰ *Id.* at 3.

¹⁰¹ See, e.g., MassHealth, *Guidelines for Medical Necessity Determination for Gender Reassignment Surgery* (Jan. 2, 2015), <https://www.mass.gov/files/documents/2017/10/18/mg-genderreassignment.pdf>.

¹⁰² See generally HEALTH INS. COMM'R, BULL. No. 2015-3, at 3 n.9 (2015) (noting that the Commissioner's analysis was "similar, in part, to that

On December 18, 2014, following the lead of the EEOC in Macy and the Office of Civil Rights opinions as set forth in the Rodriguez Letter, United States Attorney General Eric Holder announced that the Department of Justice (“DOJ”) would be taking the position in litigation that the protection of Title VII of the Civil Rights Act of 1964 extended to claims of discrimination based on an individual’s gender identity, including transgender status.¹⁰³ Attorney General Holder issued a memorandum that informed all DOJ heads and United States Attorneys that the DOJ would no longer assert that Title VII’s prohibition against discrimination based on sex excludes discrimination based on gender identity *per se*, including transgender discrimination, reversing a previous DOJ position.¹⁰⁴

Title VII makes it unlawful for employers to discriminate in the employment of an individual “because of such individual’s...sex,” among other protected characteristics.” This important shift will ensure that the protections of the Civil Rights Act of 1964 are extended to those who suffer discrimination based on gender identity, including transgender status,” said Attorney General Holder. “This will help to foster fair and consistent treatment for all claimants. And it reaffirms the Justice Department’s commitment to protecting the civil rights of all Americans.”

Under the Trump Administration, the federal government adopted a new position with respect to the application of the country’s discrimination laws to gender identity. On October 7, 2017, Attorney General Jeff Sessions issued a memorandum to all United States Attorneys and Heads of Department Components, entitled “Revised Treatment of Transgender Employment Discrimination Claims Under Title VII of the Civil Rights Act

conducted by the Division of Insurance for the Commonwealth of Massachusetts in their Bulletin 2014-03”); N.Y. State Dep’t of Fin. Serv., Ins. Circular Letter No. 7 (Dec. 11, 2014); MINN. DEP’T OF COM. & MINN. DEP’T OF HEALTH, ADMIN. BULL. No. 2015-5 (2015); NEV. DEP’T OF BUS. & INDUSTRY, DIVISION OF INS., BULL. No. 15-002, 2015 (2015); Wash. Office of the Ins. Comm’r, Letter (June 25, 2014); PENN. DEP’T OF INS., BULL. No. 16-762, (2016); MICH. DEP’T OF INS. & FIN. SERVS., BULL. No. 2016-10-INS (2016); and MD. INS. ADMIN. BULL. No. 15-33 (2015).

¹⁰³ See Press Release No. 14-1429, United States Department of Justice, Office of Press Affairs (Dec. 18, 2014) <https://www.justice.gov/opa/pr/attorney-general-holder-directs-department-include-gender-identity-under-sex-discrimination>.

¹⁰⁴ *Id.*

of 1964.”¹⁰⁵ The Attorney General stated that “[a]lthough federal law, including Title VII, provides various protections to transgender individuals, Title VII does not prohibit discrimination based on gender identity *per se*.”¹⁰⁶ In so stating, Attorney General Sessions further noted that in a December 15, 2014, memorandum, former Attorney General Holder came to the opposite conclusion, namely, that Title VII does encompass such discrimination, based on his view that Title VII prohibits employers from taking into account “sex-based considerations.”¹⁰⁷

Attorney General Sessions further stated that, upon his review of the pertinent statutory and case law, he concluded that “Title VII’s prohibition on sex discrimination encompasses discrimination between men and women but does not encompass discrimination based on gender identity *per se*, including transgender status.”¹⁰⁸ Because of his conclusion, he formally withdrew Attorney General Holder’s December 15, 2014, memorandum, and stated that the DOJ would henceforth adopt his conclusion in all pending and future matters.¹⁰⁹

Similarly, on February 22, 2017, the U.S. Departments of Education and Justice (the “Departments”) issued a “Dear Colleague Letter” that stated that the Departments were withdrawing the statements of policy and guidance reflected in two previously-issued guidance documents: the Letter to Emily Prince from James A. Ferg-Cadima, Acting Deputy Assistant Secretary for Policy, Office for Civil Rights at the Department of Education dated January 7, 2015; and the “Dear Colleague Letter” on transgender students jointly issued by the Civil Rights Division of the Department of Justice and the Department of Education dated May 13, 2016.¹¹⁰

¹⁰⁵ See Office of the Attorney General Memorandum, *Revised Treatment of Transgender Employment Discrimination Claims Under Title VII of the Civil Rights Act of 1964*, (Oct. 4, 2017) <https://www.justice.gov/ag/page/file/1006981/download>.

¹⁰⁶ *Id.* at 1.

¹⁰⁷ *Id.*

¹⁰⁸ *Id.* at 2.

¹⁰⁹ *Id.*

¹¹⁰ Sandra Battle, Acting Assistant Sec’y for Civil Rights, U.S. Dep’t of Educ. & T.E. Wheeler II, Acting Assistant Attorney General for Civil Rights, U.S. Dep’t of Just., Dear Colleague Letter (Feb. 22, 2017), <https://www2.ed.gov/about/offices/list/ocr/letters/colleague-201702-title-ix.pdf>.

The Departments noted that these guidance documents took the position that the prohibitions on discrimination “on the basis of sex” in Title IX of the Education Amendments of 1972 (“Title IX”), 20 U.S.C. § 1681 et seq., and its implementing regulations, 34 C.F.R. § 106.33, required access to sex-segregated facilities based on gender identity.¹¹¹ In the February 22, 2017 Dear Colleague letter, the Departments stated that they had decided to withdraw and rescind the above-referenced guidance documents “in order to further and more completely consider the legal issues involved,” and that “the Departments thus will not rely on the views expressed within them.”¹¹²

On October 21, 2018, the New York Times reported that the U.S. Department of Health and Human Services (“HHS”) had revealed in an internal memorandum the agency’s intention to narrow the legal definition of “sex” under Title IX.¹¹³ In the leaked memorandum that had been obtained by the New York Times, HHS urged government agencies enforcing Title IX - including the DOJ - to adopt a single, uniform definition of gender based “on a biological basis that is clear, grounded in science, objective and administrable, where “sex” meant only “a person’s status as male or female based on immutable biological traits identifiable by or before birth,” The HHS memorandum further stated that the sex “listed on a person’s birth certificate, as originally issued, shall constitute definitive proof of a person’s sex unless rebutted by reliable genetic evidence.”

On March 7, 2018, the United States Court of Appeals for the Sixth Circuit, in the case *R.G. & G.R. Harris Funeral Homes v. Equal Employment Opportunity Commission*, held that discrimination against transgender people was barred by Title VII.¹¹⁴ The Court of Appeals stated that “[i]t is analytically impossible to fire an employee based on that employee’s status as a transgender person without being motivated, at least in part, by the employee’s sex,” and “discrimination ‘because of sex’ inherently includes discrimination against employees because of a change in their sex.”¹¹⁵

¹¹¹ *See id.*, at 1.

¹¹² *See id.* at 2.

¹¹³ *See* Erica L. Green, Katie Benner & Robert Pear, ‘Transgender’ Could Be Defined Out of Existence Under Trump Administration, N.Y. TIMES, (Oct. 21, 2018), <https://www.nytimes.com/2018/10/21/us/politics/transgender-trump-administration-sex-definition.html>.

¹¹⁴ *Equal Employment Opportunity Commission v. R.G. & G.R. Harris Funeral Homes, Inc.*, 884 F.3d 560, 572 (6th Cir. 2018).

¹¹⁵ *Id.* at 575.

R.G. & G.R. Harris Funeral Homes petitioned the United States Supreme Court for a writ of certiorari from the Third Circuit's decision.¹¹⁶ On April 22, 2019, the Supreme Court granted the petition for certiorari, limited to the following question: "Whether Title VII prohibits discrimination against transgender people based on (1) their status as transgender or (2) sex stereotyping under *Price Waterhouse v. Hopkins*, 490 U. S. 228 (1989)."¹¹⁷ Oral argument before the Supreme Court is scheduled for October 8, 2019.¹¹⁸

CONCLUSION

As noted previously in this Article, the Division, in determining that the denial of coverage by health insurance companies for gender transition-related medical care including gender assignment surgery, hormone replacement therapy and other treatments based on an individual's gender identity or gender dysphoria, was sex-based discrimination prohibited under Massachusetts law. The Division had relied in part upon the Obama Administration's interpretation of existing non-discrimination law—including Title VII and Title IX—to mean that the sex-discrimination protections under the ACA also applied to transgender people. The Trump Administration has upended this interpretation and stated that it no longer views existing laws as extending "sex-discrimination" protections to transgender people. The Division's conclusions nevertheless remain intact and persuasive.

The Division, in transitioning to a new view as to what was considered to be prohibited sex discrimination in the provision of benefits under health insurance policies to transgender persons, also relied also upon the long-standing lead of the majority of federal and state courts here in Massachusetts in concluding that the term "sex discrimination" must be given a broad interpretation. This conclusion is supported by the recent action of the Massachusetts Legislature in extending additional protections to transgender persons, including the passage of legislation in 2016 to extend protections against discrimination for gender identity to any place of public

¹¹⁶ See *R.G. & G.R. Harris Funeral Homes Inc. v. Equal Employment Opportunity Commission*, SCOTUSBLOG, <https://www.scotusblog.com/case-files/cases/r-g-g-r-harris-funeral-homes-inc-v-equal-opportunity-employment-commission/>. (last visited July 19, 2019).

¹¹⁷ *Id.*

¹¹⁸ *Id.*

accommodation,¹¹⁹ and the actions of the people of Massachusetts in voting on November 6, 2018, in a ballot initiative to uphold this state law forbidding discrimination based on gender identity in public places.¹²⁰

To the extent that the Supreme Court does ultimately rule that the term “sex discrimination” under Title VII, and by extension, Title IX, does not include discrimination because of gender identity, in order to fully ensure that benefits under health insurance policies are not denied to transgender persons on account of their gender identity, the Massachusetts Legislature should consider amending Chapter 134 of the Acts of 2016¹²¹ to include specific protection for transgender persons with respect to health coverage.

More work needs to be done to protect the rights of transgender persons in seeking their rightful benefits under health insurance policies. But a good start has been made here in the Commonwealth of Massachusetts in making sure that carriers will no longer be able to discriminate against transgender individuals as they seek coverage for gender transition-related medical care.

¹¹⁹ 2016 Mass Acts. Ch. 134.

¹²⁰ See Faith Karimi & Emanuella Grinberg, *Massachusetts Voters Uphold Transgender Rights Protections*, CNN POLITICS, <https://www.cnn.com/2018/11/07/politics/massachusetts-upholds-transgender-protections/index.html> (last updated Nov. 7, 2018, 9:38 AM).

¹²¹ 2016 Mass Acts. Ch. 134.

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities related to the business. It emphasizes the need for transparency and accountability in financial reporting.

2. The second part of the document outlines the various methods and techniques used to collect and analyze data. It includes a detailed description of the experimental design and the procedures followed to ensure the reliability and validity of the results.

3. The third part of the document presents the results of the study, including a comprehensive analysis of the data collected. It discusses the key findings and their implications for the field of research.

4. The fourth part of the document provides a conclusion and discusses the limitations of the study. It also offers suggestions for future research and potential applications of the findings.

5. The final part of the document includes a list of references and a bibliography, providing a comprehensive overview of the sources used in the study.

6. The following table provides a summary of the key findings and their implications for the field of research. It includes a detailed description of the experimental design and the procedures followed to ensure the reliability and validity of the results.

7. The results of the study indicate that there is a significant positive correlation between the variables studied. This finding has important implications for the field of research and suggests that further investigation is needed to explore the underlying mechanisms.

8. In conclusion, the study has provided valuable insights into the relationship between the variables studied. The findings suggest that there is a need for further research to explore the underlying mechanisms and to develop more effective interventions.

9. The following table provides a summary of the key findings and their implications for the field of research. It includes a detailed description of the experimental design and the procedures followed to ensure the reliability and validity of the results.

10. The results of the study indicate that there is a significant positive correlation between the variables studied. This finding has important implications for the field of research and suggests that further investigation is needed to explore the underlying mechanisms.

11. In conclusion, the study has provided valuable insights into the relationship between the variables studied. The findings suggest that there is a need for further research to explore the underlying mechanisms and to develop more effective interventions.

12. The following table provides a summary of the key findings and their implications for the field of research. It includes a detailed description of the experimental design and the procedures followed to ensure the reliability and validity of the results.

13. The results of the study indicate that there is a significant positive correlation between the variables studied. This finding has important implications for the field of research and suggests that further investigation is needed to explore the underlying mechanisms.

14. In conclusion, the study has provided valuable insights into the relationship between the variables studied. The findings suggest that there is a need for further research to explore the underlying mechanisms and to develop more effective interventions.

15. The following table provides a summary of the key findings and their implications for the field of research. It includes a detailed description of the experimental design and the procedures followed to ensure the reliability and validity of the results.

PLAIN MEANING, EXTRINSIC EVIDENCE, AND AMBIGUITY: MYTH AND REALITY IN INSURANCE POLICY INTERPRETATION

KENNETH S. ABRAHAM*

Insurance coverage disputes are mostly about the correct interpretation of an insurance policy provision. But three myths confuse and confound thinking about the interpretation of insurance policies. The first myth is that an unambiguous insurance policy provision – a provision with a “plain” meaning – carries that meaning on its face. The second myth is that, if a policy provision has a plain meaning, then under the plain-meaning “rule,” sources of meaning outside the four corners of the insurance policy – sources “extrinsic” to the policy -- are not admissible to aid in interpreting the provision. The third myth is that ambiguous policy provisions are necessarily construed against the drafter, which in insurance is almost always the insurer. In reality, all three myths seriously oversimplify how interpretation takes place. The problem, however, is not that, in acting in ways that are inconsistent with the simplifying myths, the courts are undermining desirable rules by quietly following other, undesirable rules. On the contrary, we do not need to change the rules or practices that govern insurance policy interpretation; Rather, we need more clarity and a deeper understanding of the sophisticated, complex rules and practices that are actually in force and are actually applied in practice. This Article aims to provide both.

* David and Mary Harrison Distinguished Professor, University of Virginia School of Law. Thanks to Tom Baker, Kyle Logue, and Daniel Schwarcz for helpful comments.

INTRODUCTION.....	331
I. INTERPRETATION, APPLICATION, AND CONSTRUCTION...	334
A. INTERPRETATION	335
B. APPLICATION	336
C. CONSTRUCTION.....	339
II. PLAIN MEANING AND THE “WHOLE” POLICY	340
A. THREE NORMATIVE PRESUMPTIONS ABOUT THE “WHOLE” POLICY	341
1. Consistency	342
2. Coherence	343
3. Non-Redundancy.....	344
B. THE CANONS OF INTERPRETATION.....	346
III. EXTRINSIC SOURCES OF MEANING UNDER THE PLAIN- MEANING RULE.....	347
A. IMPLICIT KNOWLEDGE.....	348
B. ADJUDICATIVE FACTS SUBJECT TO JUDICIAL NOTICE.....	351
C. LEGISLATIVE FACTS.....	352
D. TAKING STOCK.....	354
IV. AMBIGUITY ABOUT AMBIGUITY: INTERPRETATION AND CONSTRUCTION OF AMBIGUOUS POLICY PROVISIONS	355
A. INTERPRETING AMBIGUOUS POLICY PROVISIONS	356
B. CONTRA PROFERENTEM: THREE LEVELS OF CONSTRUCTION AGAINST THE DRAFTER	359
1. Ambiguity by Necessity	359
2. Ambiguity Resulting from Faulty Drafting.....	363
3. Ambiguity as a Trap.....	364
CONCLUSION.....	366

INTRODUCTION

Insurance coverage disputes are mostly about the correct interpretation of an insurance policy provision.¹ The heated controversy that took place over a period of years at the American Law Institute (“ALI”) regarding the interpretation provisions of the Restatement of the Law of Liability Insurance (“RLLI”)² reflects the crucial role that the rules governing interpretation play in coverage disputes. In 2017, the ALI membership approved a rule that for several years had been included in drafts of the RLLI, permitting the introduction of extrinsic evidence, without limit, in order to determine whether a policy provision is ambiguous.³ But the following year there was an about face. The Reporters recommended and the ALI adopted an amendment embodying the plain-meaning rule, which precludes the introduction of extrinsic evidence if the meaning of a policy provision is plain on its face.⁴ This rule, and the entire RLLI, are now final.⁵

¹ A quick glance at any of the principal insurance law casebooks confirms that this is the case. *See generally*, KENNETH S. ABRAHAM & DANIEL SCHWARCZ, *INSURANCE LAW & REGULATION* (6th ed. 2015); TOM BAKER & KYLE D. LOGUE, *INSURANCE LAW AND POLICY: CASES AND MATERIALS* (3d ed. 2013); LEO P. MARTINEZ & DOUGLAS R. RICHMOND, *CASES AND MATERIALS ON INSURANCE LAW* (8th ed. 2018); JEFFREY W. STEMPEL ET AL., *PRINCIPLES OF INSURANCE LAW* (4th ed. 2011).

² RESTATEMENT OF THE LAW OF LIABILITY INSURANCE (AM. L. INST., Proposed Final Draft No. 2, 2018) [hereinafter RLLI].

³ *Id.* at § 3(2). This Section also provided for a presumption in favor of the plain meaning (if any) of the provision. This presumption had no explicit support in the case law. Rather, the minority, contextual rule, is that extrinsic evidence is admissible, but with no presumption in favor of plain meaning, to demonstrate that a provision that is unambiguous on its face contains a “latent” ambiguity. *Id.* at § 3 cmt. a; *see also* *City of Gross Pointe Park v. Michigan Mun. Liab. Pool*, 702 N.W. 2d 106, 113 (Mich. 2005); *Brown Mech. Contractors, Inc. v. Centennial Ins. Co.*, 431 So. 2d 932, 942 (Ala. 1983).

⁴ RLLI § 3.

⁵ *See, e.g.*, *Eyeblaster, Inc. v. Fed. Ins. Co.*, 613 F.3d 797, 802 (8th Cir. 2010) (stating that “[t]he plain meaning of tangible property includes computers....”).

This extended controversy, and the limitations of the restatement form, however, obscured and oversimplified both the rules governing interpretation and the process of interpretation that the rules govern.⁶ In fact, the whole controversy was emblematic of three simplifying myths that confuse and confound thinking about the interpretation of insurance policies. The first myth is that an unambiguous insurance policy provision – a provision with a “plain meaning” – carries that meaning on its face. In reality, many policy provisions are accorded a plain meaning through an active process of interpretation. Courts often do not simply *receive* a plain meaning by reading an insurance policy. Rather, they actively *construct* that single, “plain” meaning.

The second myth is that, if a policy provision has a plain meaning, then under the plain-meaning rule, sources of meaning outside the four corners of the insurance policy – sources “extrinsic” to the policy -- are not admissible to aid in interpreting the provision.⁷ In reality, important sources of meaning outside of an insurance policy may be considered, and often are considered, in interpreting policy provisions that courts then hold have an unambiguous, plain meaning.⁸

The third myth concerns ambiguous policy provisions – those that are reasonably susceptible to two different interpretations. Under the doctrine *contra proferentem* (“against the offeror” or drafter),⁹ ambiguous policy provisions are supposedly construed against the drafter, which in insurance is almost always the insurer.¹⁰ In reality, a finding of ambiguity merely authorizes the introduction of otherwise-inadmissible extrinsic evidence to aid in interpreting the ambiguous provision.¹¹ In addition, in

⁶ The RLLI did not fall prey to these myths, but the necessary requirements of both brevity and format (black-letter rules followed by concise “comments”) limited its capacity to dispel them, and foreclosed the kind of extended analysis undertaken here.

⁷ See ROBERT H. JERRY II & DOUGLAS R. RICHMOND, UNDERSTANDING INSURANCE LAW 121 (6th ed. 2018) (describing this view without endorsing it).

⁸ RLLI § 3 cmt. b.

⁹ See ABRAHAM & SCHWARCZ, *supra* note 1, at 41; RESTATEMENT (SECOND) OF CONTRACTS §206 (1981) (providing for interpretation “against the Draftsman”).

¹⁰ JERRY & RICHMOND, *supra* note 7 at 127, and at n.305 (identifying this view and citing courts adopting it).

¹¹ See RLLI § 4 cmt. b.

reality, even when the provision remains ambiguous after such extrinsic evidence is considered, the courts do not necessarily construe the provision in favor of coverage.¹²

In my view, what occurs in reality in all three respects is perfectly acceptable. The problem is not that, in acting in ways that are inconsistent with the simplifying myths, the courts are undermining desirable rules by quietly following other, undesirable rules. On the contrary, the problem is that statements the courts and commentators make often oversimplify the rules that are actually being applied, and thereby perpetuate misconceptions about the realities of insurance policy interpretation. We do not need to change the rules or practices that govern insurance policy interpretation; we need more clarity and a deeper understanding of the sophisticated, complex rules and practices that are actually in force and are actually applied in practice.¹³

This Article aims to provide both greater clarity and a deeper understanding of these rules and practices. Part I sets the stage for the analysis by distinguishing interpretation of insurance policies from both application of the policy to a claim, and construction of the policy in order to determine its legal effect. Because the plain-meaning rule applies only to interpretation, these distinctions are crucial. Part II explores the nature of insurance policy interpretation and the process of determining that policy provisions have a plain meaning by consulting the “whole” policy. The underlying insight that emerges is how active the process of arriving at a “plain” meaning can be, even when nothing “extrinsic” to an insurance policy is expressly taken into consideration.

Part III then considers the seemingly contradictory practice of expressly and openly considering certain matters that are extrinsic to the policy, even on the part of courts that follow a “strict” plain-meaning rule.

¹² See discussion *infra* Section IV.B.

¹³ The insurance law with which this Article is concerned is, in effect, insurance *contract* law. Many of the rules and concepts of insurance law are drawn straightforwardly from the law of contract interpretation. Others, however, are distinctive to insurance law, or find their most detailed elaboration and application in insurance law. In most instances there is little to be gained here from identifying in detail which rules replicate conventional contract law and which rules are distinctive to insurance law, although I will indicate important differences where appropriate, and will cite general principles of contract law when they are applicable.

Here I argue that this practice is not at all contradictory, because the matters these courts routinely consider are extrinsic to the policy but are not “evidence.” The plain-meaning rule, it turns out, is not really about plain meaning, but about which sources may be considered in determining whether a policy provision has a single reasonable meaning.

Finally, Part IV examines interpretation and construction when a policy provision is ambiguous. I show that the strong *stare decisis* effect accorded to the interpretation and construction of ambiguous policy provisions, as well as the notion that policy provisions that remain ambiguous even in the face of extrinsic evidence are automatically construed in favor of coverage, are both open to question. In short, this ambiguity about ambiguity needs examination.¹⁴

I. INTERPRETATION, APPLICATION, AND CONSTRUCTION

Courts perform three functions relating to the meaning of insurance policies. *Interpretation* is the process of determining meaning.¹⁵ *Application* is the process of determining whether, given the meaning of the relevant policy provision or provisions, a claim for coverage involving particular facts and circumstances is or is not covered. *Construction* is the process of determining the legal effect of an insurance policy,¹⁶ which may or may not

¹⁴ Over twenty years ago, I developed a conceptual framework for analyzing the factors that could influence courts’ determinations that policy provisions are ambiguous, and of the consequences of these determinations. See Kenneth S. Abraham, *A Theory of Insurance Policy Interpretation*, 95 MICH. L. REV. 531 (1996). In a sense, this Article is an extension of that framework to the particular issue of the *evidence* that is relevant to the plain-meaning/ambiguity issue, identifying factors bearing on this issue at a level of detail that my general theory did not encompass. I note at several points below where there is resonance with the earlier Article. In addition, Part IV (C) adds a factor relevant to ambiguity (ambiguity as a “trap”) that I had not recognized at that time.

¹⁵ RESTATEMENT (SECOND) OF CONTRACTS §200 (AM. L. INST. 1979); E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS §7.08 (Zachary Wolfe ed., 4th ed. 2019).

¹⁶ Edwin W. Patterson, *The Interpretation and Construction of Contracts*, 64 COLUM. L. REV. 833, 835 (1964).

coincide with its meaning as determined by interpretation. The plain-meaning rule governs interpretation, not application or construction.

A. INTERPRETATION

The plain-meaning rule prohibits considering extrinsic evidence of a policy provision's meaning when the policy provision has a plain meaning on its face.¹⁷ By its terms, then, the rule only governs interpretation – the determination of meaning. The dominance of the plain-meaning rule in insurance contrasts starkly with general contract law, where the dominant approach is to permit the introduction of extrinsic evidence to aid interpretation.¹⁸ Interpretation is a matter for the court, unless it depends on factually-disputed issues.¹⁹

It is sometimes said that a policy provision may be ambiguous in a particular context but unambiguous in another context.²⁰ The logical implication of such statements is that the provision has a plain meaning in one context but not a plain meaning in another context.²¹ For example, CGL insurance policies typically contain a provision that excludes coverage of liability for damage to property in the “care, custody or control” of the insured.²² It may be indisputable that an insured holding an item of personal

¹⁷ See RLLI §3.

¹⁸ There is a plain-meaning rule in the general law of contracts, but many jurisdictions have rejected it, even while adopting the plain-meaning rule for insurance contracts. *Id.* at cmt a; JOSEPH M. PERILLO, CONTRACTS §3.10 at 136 (7th ed. 2014). The “modern” view in general contract law expressly permits the introduction of many forms of extrinsic evidence regardless of ambiguity. RESTATEMENT (SECOND) OF CONTRACTS §202 (AM. LAW INST. 1979) (providing that contracts are to be interpreted “in light of all the circumstances”) and §202(4)-(5) (providing that interpretation is to take place as consistent with relevant course of performance or dealing and usage of trade). In addition, evidence of prior negotiations is admissible to establish the meaning of the contract under many conditions. *Id.* at §214(c) It would be only a slight exaggeration, therefore, to say that there is no plain-meaning rule under the RESTATEMENT (SECOND) OF CONTRACTS.

¹⁹ See RLLI §2(2); RESTATEMENT (SECOND) OF CONTRACTS, *supra* note 9, at §212(2).

²⁰ See, e.g., ABRAHAM & SCHWARCZ, *supra* note 1, at 45.

²¹ See RLLI §3 cmt. f.

²² ABRAHAM & SCHWARCZ, *supra* note 1, at 443.

property in his arms has the property in his “care, custody or control.” The provision therefore has a plain meaning in this context. On the other hand, whether a parcel that has been delivered and left on the doorstep of the insured is in the insured’s “care, custody or control” may be debatable. The provision arguably does not have a plain meaning in this context.

A different way of making this point would be to say that the “care, custody or control” exclusion has a plain meaning “as applied” to the claim involving personal property held in the insured’s arms, but is ambiguous “as applied” to the parcel left on the insured’s doorstep. There is nothing wrong with this alternative formulation in itself, but it does risk confusing the process of interpretation with the process of application. Under the plain-meaning rule, extrinsic evidence would not be admissible to interpret – to determine the meaning of – the “care, custody or control” exclusion, whether in the abstract or “as applied” to either of these claims for coverage. However, as indicated next, extrinsic evidence about either claim would be admissible to aid in the application of an interpretation made under the plain-meaning rule to a particular claim.

B. APPLICATION

The plain-meaning rule does not preclude the introduction of all extrinsic evidence. There is no prohibition on the admission of extrinsic evidence in order to apply a policy provision to a claim for coverage. For example, without evidence of the facts associated with a claim, the policy could not be applied to a claim. In my earlier hypothetical, evidence of the number of steps from the walkway up to the door of the insured’s home would be admissible, because this evidence would not bear on the meaning of the “care, custody or control” exclusion. Thus, extrinsic evidence is not admissible regarding the major premise of the interpretive syllogism (i.e., “custody means...”), whereas extrinsic evidence regarding the minor premise (“a parcel was left on the insured’s doorstep under the following conditions...”) is admissible. And this kind of evidence – that the insured suffered a loss, the conditions under which the loss occurred, the amount of the loss, and so forth – is necessarily extrinsic to the policy.²³

It is undoubtedly true that, in the course of applying a policy provision to the facts of a claim, interpretation sometimes must occur. But this is not inconsistent with the distinction between interpretation and application. Interpretation does not have to be completed before application begins, in order still to constitute interpretation. Policy provisions may seem

²³ See RLLI § 2 cmt. f.

to have a meaning in the abstract or in general that must be adjusted in the course of application, or provisions may have a meaning that is too abstract or too general to determine whether a claim is covered, until the particular facts and details of a claim are known. The interpreting court can take these claim-related facts into account in determining the meaning of the policy provision to be applied to these facts, or in sharpening that meaning in light of these particular facts. This does not violate the plain-meaning rule. What matters is that the facts of the claim, and any other extrinsic evidence that is admitted, be only a *predicate* to determining meaning, not a *source* of meaning.

Thus, the process of applying the meaning of a policy provision to the facts of a claim cannot always be altogether divorced from the *act* of interpretation, but it can be divorced from the *concept and function* of interpretation. An interpretation, standing alone, is like a rule – its meaning has a level of generality that is not necessarily self-applying, any more than legal rules are always self-applying. And just as applying a legal rule to a set of facts sheds light on the meaning of the rule, so the effort to apply the interpretation of a policy provision to a claim may sometimes shed additional light on the meaning of the provision.

Applying an interpretation to a claim may involve an implicit act of mini-, or concrete, interpretation. For example, if an auto liability insurance policy covers liability for injury “arising out of the use” of an auto, determination that “use” means to drive or otherwise employ would not automatically resolve the questions whether “use” includes throwing a firecracker out the window of a parked car.²⁴ Application of the term “use” to this set of facts requires what amounts to further interpretation in this concrete circumstance. The interpretation may be only implicit in the result, or the court may explain why these facts do or do not constitute a “use,” thus expressly interpreting that term in this particular context. The combination of a series of applications to similar claims involving slightly different facts may produce what amounts to a more detailed interpretation of the term “use.” But the facts of the claim in these situations serve only as a predicate to, not a source of, the interpretation.

Similarly, in *Stone Container Corp. v. Hartford Steam Boiler Inspection and Ins. Co.*,²⁵ a boiler and machinery insurance policy excluded coverage of losses caused by explosion, with an exception for losses caused

²⁴ See *Farm Bureau Mut. Ins. Co. v. Evans*, 637 P.2d 491, 494 (Kan. Ct. App. 1981).

²⁵ 165 F.3d 1157 (7th Cir. 1999).

by explosion of an “object of a kind described below.” Included in the list objects described below were “(1) Steam boiler” and “(2) Electric steam generator.” The policyholder made a claim for coverage of a loss resulting from the explosion of a pulp digester. The court held that these policy provisions were unambiguous, and that a pulp digester did not satisfy the requirement that it be “of a kind” with the listed objects, because (although a pulp digester was “closest to a steam boiler”), a steam boiler creates steam by boiling water, whereas the steam in a pulp digester is generated outside and then fed into the digester.²⁶

Although the court’s opinion (by Posner, J.) did not recount the court’s thought process, it seems highly likely that the court did not first determine in the abstract what kinds of objects were “of a kind” with those listed, and only after determining what characteristics these objects had in common, then turn to the facts of the claim to identify the characteristics of pulp digesters. Rather, in all probability the court attributed a provisional meaning to the policy provision, looked at the record evidence and thought about the characteristics of pulp digesters, reflected again about the meaning of the policy provision, and through this process of provisional interpretation and provisional application, arrived at a conclusion that applied the now better-understood meaning of the provision to the claim for coverage of losses caused by the pulp digester’s explosion. Interpretation of the provision and application of the meaning arrived at through interpretation to the pulp digester involved an iterative, or reflexive, process.

When both interpretation and application are matters for the court, then all this is mainly a matter of nomenclature, for both are then subject to appellate review. In contrast, when interpretation is for the court but application, even in the absence of a dispute about the empirical facts, is a question of fact (as it is in a minority of jurisdictions²⁷), then applications, including the concrete interpretations that follow from application, may vary from case to case, even when the relevant facts are identical. If application is a question of fact, then some explosions of pulp digesters will be covered, and some will not be covered, by the same insurance policy that was at issue in *Stone Container*, depending on the application of the policy language by the trier of fact in each individual case.

This seems undesirable, in light of the fact that standard-form insurance policies should provide the same coverage to identically-situated insureds. The dominant, and I think preferable, approach, is therefore for

²⁶ *Id.* at 1160.

²⁷ *See* RLLI § 2 cmt. f.

both interpretation and application to be matters for the court. When application depends on the resolution of a purely empirical dispute (e.g., whether this pulp digester generated its own steam), the need for a finding of fact can be satisfied by asking the jury to bring in a special verdict, to be followed by the court's applying its interpretation to the factual findings contained in the that verdict.

The central point, however, is that the facts of a claim are not, and may not be permitted to be, a source of meaning. Rather, these facts may stimulate and focus the court's thinking about the meaning of the relevant policy provision on its face, and about the proper application of this meaning to the claim. The two functions, interpretation and application, are logically and conceptually distinct, even when they occur in an iterative sequence and the facts of the claim help to inform the court's thinking. Interpretation is logically prior to application, even when the two are temporally mixed.

C. CONSTRUCTION

Interpretation must also be distinguished from construction, which is the process of determining the legal effect of a policy provision, or any other contract.²⁸ A policy provision can have one meaning (or more than one) but a different legal effect. *Contra proferentem* is a rule of construction, addressing the legal effect of ambiguous policy language. Similarly, the rule that policy language that affirmatively provides coverage should be construed broadly, and language (such as an exclusion) restricting coverage should be construed narrowly is, as it states, a rule of construction.²⁹ Whether this rule is anything other than an application of *contra proferentem* is not entirely clear, since some courts appear not to treat it this way, but that question is not pertinent here. And the invalidation of a policy provision on the ground that it violates public policy—for example, by covering liability for punitive damages³⁰—is likewise an act of construction rather than interpretation.

Construction is sometimes camouflaged as interpretation, either unintentionally or deliberately, in order to obscure the extent of a court's

²⁸ *Id.* §2 cmt. (g); FARNSWORTH, *supra* note 15, § 7.08.

²⁹ See JERRY & RICHMOND, *supra* note 7, at 127.

³⁰ See generally *Hartford Cas. Ins. Co. v. Powell*, 19 F. Supp. 2d 678 (N.D. Tex. 1998).

lawmaking.³¹ This kind of conflation of the two functions can lead to uncertainty about what rules govern interpretation. A prominent example involves the letter sent to insurance commissioners in the early 1970s by insurance industry rating bureaus seeking state regulatory approval of the incorporation of a qualified “pollution exclusion” into the standard-form Comprehensive General Liability (CGL) insurance policy of the period. The proposed provision excluded coverage of liability for bodily injury or property damage caused by the discharge of pollutants, but contained an exception for discharges that were “sudden and accidental.” Some years after the exclusion was approved, policyholders contended in coverage disputes that, among other things, this letter’s assertions about the meaning and effect of the exception to the exclusion were misleading, and that insurers should therefore be estopped to assert that the term “sudden” had a temporal component.³²

This is best understood as an argument about the proper construction of the term “sudden,” not about the interpretation of that term. The argument was, in effect, that even if the plain meaning of “sudden” within the four corners of the policy included a temporal component, the policy should not be enforced to give the provision its plain, temporal meaning. Because construction in such instances is a judicial intervention upsetting the meaning of a policy provision, courts may sometimes understandably be reluctant to acknowledge that they are engaged in construction rather than interpretation. Extrinsic evidence, such as the letter to insurance commissioners, should be admissible as relevant to construction, even when it is not admissible for purposes of interpretation. Confusing or conflating the two processes risks obfuscating the rule regarding the evidence that may be considered when interpretation, and not construction, occurs.

II. PLAIN MEANING AND THE “WHOLE” POLICY

Courts following the plain-meaning rule do not merely stare at the words of a policy provision in order to determine whether the provision is ambiguous, or to determine what the provision means once they conclude that it is not ambiguous. Thus, plain meaning is not a self-evident fact. The

³¹ See, e.g., FARNSWORTH, *supra* note 15, §7.08, at 7-75 (stating that courts “have more often ignored [this distinction] by characterizing the process of ‘construction’ as that of ‘interpretation’ in order to obscure the extent of their control over private agreement”).

³² For discussion of this multi-year episode, see *American States Ins. v. Koloms*, 687 N.E. 2d 72, 79-82 (Ill. 1997).

conclusion that a policy provision has a plain meaning is itself the result of an interpretive process that is not always simple or one-dimensional. Typically, a plain meaning does not find the court. Rather, the court finds a plain meaning. It turns out that there is a lot more to plain meaning, and to the plain-meaning rule, than meets the eye.

A frequent formulation of the prohibition on the admission of extrinsic evidence is that the court must stay within the “four corners” of the policy in determining the meaning of the disputed provision.³³ Of course, if in doing so the court determines that the provision is not ambiguous, it has simultaneously determined the plain meaning of the provision, because an unambiguous policy provision is one that has only a single reasonable meaning.

The four-corners limitation reveals little, however, about how active the process of interpretation that is confined in this way actually is permitted to be, and often is. The material within the four corners of the policy is, obviously, the whole policy. The plain-meaning rule therefore not only permits consulting the whole policy to determine the meaning of a particular provision; it would be imprudent not to consult the whole policy in doing so.

Just as courts do not stare at a policy provision in order to determine if it has a plain meaning, they do not merely read the whole policy to help determine the meaning of a particular provision. First, a set of normative presumptions about how the “whole” insurance policy has been constructed and functions serve to guide interpretation of policy provisions whose meaning might otherwise be in doubt. Second, the canons of interpretation – which surprisingly have not been recognized to be directly about the relevance of the “whole” policy or contract – often provide strong direction about the significance of other provisions or terms in the policy for a disputed provision’s meaning.

A. THREE NORMATIVE PRESUMPTIONS ABOUT THE “WHOLE” POLICY

Both the general injunction to read the policy as a whole³⁴ and the canons of interpretation do more than confirm that the meaning of one policy

³³ See, e.g., JERRY & RICHMOND, *supra* note 7, at 121.

³⁴ The notion that the whole contract is to be considered is a principle of both general contract law and insurance law. See RLLI § 3 cmt. g; RESTATEMENT (SECOND) OF CONTRACTS §202(2); National Union Fire Ins.

provision can shed light on the meaning of another provision. These principles of interpretation also reflect the notion that an insurance policy is a functioning mechanism, containing different parts that work together. Recognizing the functional quality of insurance policies reveals three features of the “whole” policy that are reflected in the courts’ approach to the interpretation of insurance policies.

These are the soft presumptions of consistency, coherence, and non-redundancy. The courts often do not state that they are following or invoking these presumptions, in part because the presumptions are so fundamental as to be almost transparent. But the courts follow them, nonetheless. This is because an insurance policy is not only a contract, but a communication of the terms of the contract to the parties and to the courts. The courts assume that, other things being equal, the parties have attempted not to contradict or unnecessarily repeat themselves, because these are features of effective and rational communication.

Nonetheless, because language is an imperfect instrument of communication, and the drafters of insurance policies sometimes imperfectly employ this imperfect instrument, the provisions contained in insurance policies are not always consistent, coherent, and non-redundant. For this reason, in practice the presumption that insurance policies have these characteristics are soft presumptions only, working propositions with an “other things being equal” quality to them.

1. Consistency

The strongest presumption is that policy provisions do not contradict each other. There is obviously something of a continuum running from complete consistency among policy provisions, to mere coherence, to lack of coherence, and finally to outright contradiction. There is at least a qualitative difference between an interpretation that avoids outright contradiction and one that goes further, by ensuring coherence among policy provisions. A policy provision may be out of keeping with the remainder of an insurance policy without directly contradicting another provision. In this situation a provision that does not cohere with the remainder of the policy could nonetheless be interpreted without the other provisions failing to function. But outright contradiction would render at least one of two inconsistent provisions inoperative. Insurers that draft standard-form

Co. v. Fed. Ins. Co., 213 F. Supp. 3d 1333, 1340 (D. Colo. 2016); Progressive Cas. Ins. Co. v. Fed. Deposit Ins. Co., 80 F. Supp. 3d 923, 946 (N.D. Iowa 2015).

policies do not intend to include contradictory provisions in their policies, nor would policyholders intend to purchase a policy containing contradictory provisions.

It is for this reason, I think, that cases involving outright contradiction are rare.³⁵ When there is a real contradiction, the courts tend to hold that the conflict between two unambiguous, contradictory provisions creates an ambiguity. For example, in *Rusthoven v. Commercial Standard Ins. Co.*,³⁶ a policy contained two contradictory Endorsements. The court held that the contradiction created ambiguity, and interpreted the policy against its drafter, the insurer.³⁷ I have not found any case in which a party argued in favor of an interpretation that would blatantly contradict the plain meaning of another provision without asserting that the result was ambiguous policy language. Rather, the argument made in such situations is that the policy is ambiguous.

The virtually complete absence of cases in which the plain meaning of one policy provision is given precedence over the plain meaning of another provision that contradicts it, without a holding of ambiguity, is evidence of how powerful the presumption of consistency is. No one argues for an interpretation that would contradict the plain meaning of another policy provision, both because insurers try mightily not to draft contradictory language, and because the presumption of consistency is so strong. Reconciling apparent inconsistency is the name of the game.

2. Coherence

A second principle that follows from the notion that an insurance policy is a functional vehicle of communication that should be read as a whole is that a policy is likely to be coherent.³⁸ This means that, when a provision can be read to cohere with the other provisions in the policy, it

³⁵ See, e.g., *PBM Nutritionals, LLC v. Lexington Ins. Co.*, 724 S.E.2d 707, 712-13 (Va. 2012) (rejecting contention that two policy provisions were in conflict).

³⁶ 387 N.W.2d 642 (Minn. 1986).

³⁷ *Id.* at 644-45.

³⁸ *But see In re SRC Holding Corp.*, 545 F.3d 661, 668 (8th Cir. 2008) (stating that “whether policy coverage ‘makes sense’ as a business matter is largely irrelevant....”).

should be read to cohere.³⁹ Presuming coherence is not the same as interpreting to avoid direct contradiction. There may be no literal or actual contradiction or inconsistency between one provision and the entire remainder of the policy, but one interpretation of a provision might nonetheless be out of keeping with the remainder.⁴⁰

For example, in *Liristis v. American Family Ins. Co.*,⁴¹ the insureds' home was contaminated by mold, apparently as a result of water used to extinguish a fire at the property. Their homeowners insurer denied coverage, relying on an exclusion providing that the policy did not cover "loss to the property...resulting directly or indirectly or *caused* by any one or more of the following...c. smog, rust, corrosion, frost condensation, mold wet or dry rot..."⁴² The court held that the loss was not excluded, because the mold contamination was not a cause of loss, but the loss itself. It would not have contradicted this language to hold that the mold contamination was excluded. It would have been plausible to hold that mold contamination resulted from mold. But the policy language did distinguish between the cause of a loss and the loss itself. The court's holding in effect took the position that the exclusion in question should be interpreted so as to be coherent with the policy's distinction between causes of loss, on the one hand, and loss itself, on the other hand.

3. Non-redundancy

The softest presumption arising out of the injunction that an insurance policy should be read as a whole is that policy language is not redundant – that every provision in a policy has an independent meaning.⁴³ Every

³⁹ The principle that every provision should be accorded some meaning is an aspect of this notion. *See, e.g., Bedford Internet Office Space, LLC v. Travelers Ins. Co.*, 41 F. Supp. 3d 535, 547 (N.D. Tex. 2014) (interpreting a policy so as to avoid rendering a policy provision of no effect).

⁴⁰ *See, e.g., S. Tr. Ins. Co. v. Phillips*, 474 S.W.3d 660, 669 (Tenn. Ct. App. 2015) (holding that policy's consistently drawing a distinction between fire, on the one hand, and vandalism and malicious mischief, on the other hand, was significant).

⁴¹ 61 P.3d 22 (Ariz. 2002).

⁴² *Id.* at 24 (emphasis added).

⁴³ This is a general principle of contract law interpretation. *See* RESTATEMENT (SECOND) OF CONTRACTS § 229 (AM. LAW INST. 1981); FARNSWORTH, *supra* note 15, §7.13.

provision, that is, is presumed to do work and to be a necessary part of the policy.⁴⁴

The reason this presumption is so soft is that it is in tension with a counter-tendency that also is sometimes exhibited by the drafters of insurance policies. This is the practice of sometimes including duplicative provisions in order to ensure or emphasize the importance of a limitation on coverage. This is the familiar “belt and suspenders” approach that is employed in the drafting of many legal documents.⁴⁵ Because the drafters want to avoid uncertainty, as well as the disputes and litigation that uncertainty may yield, they sometimes include duplicative provisions out of an excess of caution.

For example, in *TMW Enterprises v. Federal Ins. Co.*, the policyholder argued that the insurer’s interpretation of an all-risk property insurance policy rendered the “ensuing loss” in the policy “superfluous, empty words with no independent function.”⁴⁶ The court responded with an interpretation that gave the clause an independent meaning, but then continued: “But even if we choose to label this type of drafting a form of redundancy, which we do not think it is, that label surely is not a fatal one when it comes to insurance contracts, where *redundancies abound*.”⁴⁷ That phrase seems to have resonated with subsequent courts. It has since been cited in a number of opinions addressing putative redundancies in insurance policies.⁴⁸

⁴⁴ See, e.g., *Phoenix Ins. Co. v. Infogroup, Inc.*, 147 F. Supp. 3d 815, 822 (S.D. Iowa 2015) (stating that “The Court interprets undefined words in the context of the policy as a whole, and avoids interpreting the policy in such a way as to render parts of a contract ‘surplusage’”); *Northrup Grumman Corp. v. Factory Mut. Ins. Co.*, 805 F. Supp. 2d 945, 951 (C.D. Cal. 2011) (applying the principle of non-redundancy).

⁴⁵ See, e.g., *Cactus Ave., LLC v. Fid. & Guar. Ins. Co.*, No. E051787, 2012 WL 649966, at * 4 (Cal. Ct. App. Feb. 28, 2012) (indicating that “Here, similarly, an insurer could understandably want to take a ‘belt-and-suspenders’ approach and thus exclude losses under the seepage exclusion, the water exclusion, or both”); *In re SRC Holding Corp.*, 545 F.3d 661, 670 (8th Cir. 2008) (“nothing prevents the parties from using a belt and suspenders approach in drafting the exclusions, in order to be doubly sure”).

⁴⁶ *TMW Enters. v. Fed. Ins. Co.*, 619 F.3d 574, 577 (6th Cir. 2010).

⁴⁷ *Id.* (emphasis added).

⁴⁸ See, e.g., *Ardente v. Standard Fire Ins. Co.*, 744 F.3d 815, 819 (1st Cir.

B. The Canons of Interpretation

The function of a series of “canons” of interpretation, well known in contract law generally, is to identify a number of commonly-occurring relationships between or among provisions, and to specify the significance of these relationships. A number of the canons of interpretation are in fact directed primarily at the implications of the provisions in the remainder of the policy for the meaning of a disputed provision.⁴⁹

For example, *expressio unius est exclusio alterius* directs that the expression of one thing should be considered the exclusion of another thing that is not expressed.⁵⁰ If a liability insurance policy provides that it covers liability for “damages,” it is a fair inference that the absence of any mention of liability for “restitution” implies that the latter is not covered.⁵¹

Similarly, the canon *ejusdem generis* indicates that where general language is accompanied by a list of examples, the general language is to be interpreted as referring to things of the same kind as are listed.⁵² Thus, an exclusion of coverage of liability for injury or damage caused by “war,” including “undeclared or civil war,” as well as “warlike action by a military force,” and “insurrection, rebellion, [and] revolution” implies that the term “war” does not include terrorism.⁵³

Finally, under the canon *noscitur a sociis*, the meaning of a word is to be understood by reference to the meaning of the words around it.⁵⁴ According to this canon, an exclusion referring to the “release” of pollutants, as part of a list referring to the “discharge, dispersal, release or escape” of

2014); U.S. Liab. Ins. Co. v. Benchmark Constitution Servs., 797 F.3d 116, 123 (1st Cir. 2015).

⁴⁹ For discussion of the canons, see JOHN EDWARD MURRAY, JR., MURRAY ON CONTRACTS 470-71 (5th ed. 2011); 2 E. ALLEN FARNSWORTH, FARNSWORTH ON CONTRACTS 293-95 (3d ed. 2004).

⁵⁰ See MURRAY, *supra* note 49, at 471; Farnsworth, *supra* note 49, §7.13.

⁵¹ ABRAHAM & SCHWARCZ, *supra* note 1, at 466. See also Old Republic Ins. Co. v. Stratford Ins. Co., 777 F.3d 74, 81 (1st Cir. 2015) (applying, in effect, *expressio unius*, by noting that the policy made reference to one feature of the insured’s business and notably did not refer to another feature).

⁵² See MURRAY, *supra* note 49, at 471; FARNSWORTH, *supra* note 49, §7.13.

⁵³ See ABRAHAM & SCHWARCZ, *supra* note 1, at 469.

⁵⁴ See MURRAY, *supra* note 49, at 470.

pollutants, for example, would be interpreted to be listing means by which pollutants may be freed *from* confinement, and not to the deposit of pollutants *into* a place of confinement.⁵⁵

In each of these situations, the application of a canon about the significance of other policy language helped to render unambiguous a policy provision that might otherwise be regarded as reasonably susceptible to two different interpretations. The meaning of “damages” was clarified by virtue of the absence of any reference to “restitution” in liability insurance policy; the term “war” was interpreted not to include “terrorism” because of the examples of “war” included in the policy; and the meaning of “release” was interpreted by reference to the list of similar terms surrounding it. In all three situations the canons were, in effect, applications of the more general injunction that the interpretation of a disputed policy provision should not occur in isolation from the rest of the policy. Rather, the policy is to be read as a whole.⁵⁶

The lesson of my examination of the injunction to read the policy as a whole is that identification of a policy provision’s plain meaning is often an active process. The conclusion that a policy provision has a plain meaning means only that, based on the sources of meaning that may be consulted, the provision has a single reasonable meaning that must be deemed “plain” by virtue of the process of interpreting it. The plain-meaning rule is not about “plain” meaning, but about the sources that may be consulted to determine whether a policy provision has only one reasonable meaning. These observations about the complexity of the process of determining plain meaning are rendered all the more forceful once the matter that lies outside the four corners of the policy, but still may be considered under the plain-rule, are brought into view.

III. EXTRINSIC SOURCES OF MEANING UNDER THE PLAIN-MEANING RULE

Even under the plain-meaning rule, courts routinely and expressly consult certain sources of meaning that are outside the four corners of the insurance policy. These include facts that are so fundamental that they do not

⁵⁵ See, e.g., *Bd. of Regents v. Royal Ins. Co. of Amer.*, 517 N.W. 2d 888, 891 (Minn. 1994).

⁵⁶ See MURRAY, *supra* note 49, at 470; FARNSWORTH, *supra* note 49, §7.13.

even need to be articulated as sources of meaning; the “purpose” of a form of insurance coverage or a particular policy provision; dictionary definitions; other judicial decisions, statutes and regulations; and secondary legal sources.⁵⁷

The RLLI notes in a comment that, although the majority of courts follow the plain-meaning rule, these courts sometimes differ about which sources outside an insurance policy may be considered in interpreting an unambiguous policy provision. In this sense, the RLLI suggests, it might be said that there is not a “single” plain-meaning rule. The differences, however, are minimal. A few plain-meaning courts have taken the position that evidence of custom, practice, and usage may be considered even in interpreting an unambiguous policy provision.⁵⁸ These courts seem to treat such matters as “legislative facts” of the sort I discuss below in Section C. But for the most part, courts that subscribe to the plain-meaning rule do not consider custom, practice, and usage when interpreting unambiguous policy provisions.

How is it that certain sources of meaning outside of the policy can be considered, notwithstanding the prohibition against considering extrinsic evidence? The answer is that each of the above sources of meaning may be considered, despite the fact that they lie outside the four corners of the policy, because they are form of “implicit” knowledge without which judicial reasoning could not take place; because they are facts subject to judicial notice; or because they are “legislative” facts that are not subject to the rules of evidence.

A. IMPLICIT KNOWLEDGE

Insurance policy provisions are not self-defining. Modern contract theory has long recognized that a particular interpretation may be simple, straightforward, and incontestable, but that it is an interpretation nevertheless, even when it is the only reasonable interpretation.⁵⁹ This is because the reader of a contract, such as an insurance policy, including the

⁵⁷ See RLLI § 3 cmt. b.

⁵⁸ See, e.g., *Sunbeam Corp. v. Liberty Mut. Ins. Co.*, 781 A.2d 1189, 1193 (Pa. 2001); *Auto-Owners Ins. Co. v. Anderson*, 756 So. 2d 29, 36 (Fl. 2000); *Beazley Ins. Co., Inc. v. Ace American Ins. Co.*, 880 F. 3d 64, 69 (2d Cir. 2018); RLLI § 3 cmt. c.

⁵⁹ See FARNSWORTH, *supra* note 15, at §7.11 (arguing that it is questionable whether a word has a meaning at all when divorced from the circumstances).

judicial reader, always encounters contract language in a context, and always brings to bear what he or she already knows or supposes to be the relevant context when understanding – and therefore when interpreting – the meaning of that language. For the legal reader, this includes background understandings of the legal and insurance market contexts in which an insurance policy operates.

Sometimes courts expressly articulate the context in which the interpretive task is situated. But often that context is so transparent to courts, and courts expect that context to be so transparent to the legal readers to whom the court's opinion is mainly addressed, that it does not occur to the court that making explicit what is implicit in the court's reasoning is necessary. But logically, this context – which lies outside the four corners of the insurance policy – is a source that contributes to the meaning of the policy provision being interpreted. As James Bradley Thayer noted in the first modern treatise on evidence over a century ago, “In conducting a process of judicial reasoning, as of other reasoning, not a step can be taken without assuming something which has not been proved; and the capacity to do this with competent judgement and efficiency, is imputed to judges and juries as part of their necessary mental outfit.”⁶⁰

I call these assumptions and sources of meaning “implicit knowledge.” For example, many liability insurance policies cover liability incurred “because of” bodily injury or property damage.⁶¹ To the best of my knowledge, no court has ever held this phrase to be ambiguous. Nonetheless, to understand what the words mean, it is necessary to know that damages awarded in tort cases alleging bodily injury or property damage may include losses that are the consequence of the injury damage in question, such as medical expenses for treating bodily injury, or profits lost because of damage to property. Liability for these kinds of consequential losses is imposed “because of” bodily injury or property damage, even if it is not “for” such injury or damage.

Because all courts know this, the words “because of” in liability insurance policies seem in most cases to carry their meaning “on their face,” without needing any interpretation. But that only appears to be the case. It is the legal and insurance context in which the words “because of” are used in

⁶⁰ JAMES BRADLEY THAYER, *PRELIMINARY TREATISE ON EVIDENCE* 279-80 (1898).

⁶¹ *See, e.g.,* ABRAHAM & SCHWARCZ, *supra* note 1, at 439 (CGL insurance policy).

liability insurance policies that renders them subject to only one reasonable interpretation. However, when a claim for coverage of an unconventional form of liability arises – for example, when the party seeking to recover damages from the policyholder that are the consequence of bodily injury is not the same party who suffered bodily injury – then the courts must become more explicit what these words mean.⁶²

Sometimes the legal and insurance context that informs the court’s interpretation is more complex, but still “goes without saying.” For example, in *Federal Ins. Co. v. Raytheon Co.*, the policyholder claimed coverage under its Directors & Officers (D&O) liability insurance policy for liability incurred in an ERISA action; It had earlier been a defendant in a different securities law suit. The D&O insurers contended that coverage of liability in the second suit was excluded under a “pending and prior litigation” exclusion in their policies. The exclusion applied to claims against the insured “based upon, arising from, or in consequence of any demand, suit or other proceeding” pending against the insured prior to a specified date. The court held that the exclusion applied if the “allegations in the second complaint find substantial support in the first complaint,” and concluded that they did.⁶³

In order to arrive at this interpretation, however, the court had to have an understanding the complexity of ERISA and securities law suits generally, and the consequent detail that complaints typically contain, including the standard allegations regarding jurisdiction, identity of the parties, and remedies, that all such complaints make. These boilerplate allegations would have no bearing on whether the complaints substantially overlapped, since they are allegations that most complaints in complex civil suits would contain. If these sorts of allegations had been relevant to the application of the exclusion, however, then the overlap between the two complaints would have appeared to be far more substantial than it actually was – indeed, there probably would have been no issue even worth litigating. The court did not articulate any of this background context, and it did not need to do so. These facts about complex civil litigation were implicit knowledge that were one of the sources of the meaning of the unambiguous “prior and pending litigation” exclusion.

In short, it is inevitable and completely proper for courts to rely on facts outside the four corners of an insurance policy that are necessary to an

⁶² See, e.g., *Cincinnati Ins. Co. v. H.D. Smith*, 829 F.3d 771, 774-45 (7th Cir. 2017) (holding that state of West Virginia’s costs for addressing opiate addiction epidemic were incurred “because of” bodily injury, despite the fact that West Virginia itself suffered no bodily injury).

⁶³ See *Fed. Ins. Co. v. Raytheon Co.*, 426 F.3d 491 (1st Cir. 2005).

understanding of the meaning of policy provisions that the courts must interpret. Ordinarily such facts would not be in dispute if they were made explicit. But the meaning of the policy provision being interpreted might be different if the facts were otherwise. The facts may “go without saying,” in both senses of this phrase, but they are nonetheless sources of meaning extrinsic to the language of the insurance policy itself that are routinely sources relevant to the meaning of insurance policy provisions.

B. ADJUDICATIVE FACTS SUBJECT TO JUDICIAL NOTICE

More than fifty years ago, Kenneth Culp Davis distinguished between “adjudicative” and “legislative” facts. The former are facts that pertain specifically to the facts of a particular case, whereas the latter pertain to legal reasoning or the formation of legal principles.⁶⁴ Judicial notice is the process by which a court recognizes as true an adjudicative fact so well known and indisputable that it does not need to be formally introduced as evidence.⁶⁵

A court may take judicial notice of adjudicative facts if the facts are “generally known” or “can be accurately and readily determined from sources whose accuracy cannot reasonably be questioned.”⁶⁶ Certainly dictionary definitions of terms used in an insurance policy fall into this category. The same is true of such matters as the fact that the most direct route of flight between New York and Miami is partly over water more than three miles outside the territorial limits of the United States, a fact that figured in a well-known case involving the interpretation of policy language requiring that crashes occur “within” the United States.⁶⁷ Adjudicative facts of this sort lie outside the four corners of an insurance policy. But courts may

⁶⁴ Kenneth Culp Davis, *An Approach to Problems of Evidence in the Administrative Process*, 55 HARV. L. REV. 364, 404-07 (1942).

⁶⁵ Judge Posner also has observed that some “information tends to fall somewhere between facts that require adversary procedure to determine and facts of which a court can take judicial notice,” candidly acknowledging that “judges and their law clerks often conduct research on cases, and it is not always research confined to pure issues of law, without disclosure to the parties.” *Rowe v. Gibson*, 798 F.3d 622, 628 (7th Cir. 2015).

⁶⁶ FED. R. EVID. 201(b).

⁶⁷ See *Vargas v. Ins. Co. of N. Am.*, 651 F.2d 838 (1981).

take judicial notice of such facts and rely on them as sources of the meaning of policy provisions.

C. LEGISLATIVE FACTS

In addition, the rules of evidence do not preclude courts from considering what have been called "legislative facts," to distinguish them from adjudicative facts that pertain to the particular dispute.⁶⁸ A classic example of a legislative fact is the proposition that testimony by one spouse against another in a criminal proceeding would undermine most any marriage.⁶⁹ The purpose of a particular form of insurance, or of a particular standard-form policy provision, would generally fall into this category as well.

In contrast to adjudicative facts that are subject to judicial notice, which must effectively be indisputable, legislative facts need not satisfy this test:

[J]udge-made law would stop growing if judges, in thinking about law and policy, were forbidden to take into account facts they believe, as distinguished from facts which are "clearly...within the domain of the indisputable." Facts most needed in thinking about difficult problems of law and policy have a way of being outside the domain of the clearly indisputable.⁷⁰

Rather, a court's authority to consider legislative facts "renders inappropriate any limitation in the form of indisputability, any formal requirements of notice other than those already inherent in affording opportunity to hear and be heard and exchanging briefs, and any requirement of formal findings at any level."⁷¹

There are numerous cases in which the courts consider legislative facts outside the four corners of the insurance policy without ever holding

⁶⁸ FED. R. EVID. 201 Advisory Committee's note (a); Davis, *supra* note 64.

⁶⁹ FED. R. EVID. 201 Advisory Committee's note (a) (citing *Hawkins v. United States*, 358 U.S. 74, 79 (1958)).

⁷⁰ *Id.* (quoting Kenneth Culp Davis, *A System of Judicial Notice Based on Fairness and Convenience*, in ROSCOE POUND ET AL., *PERSPECTIVES OF LAW* 69, 82 (1964)).

⁷¹ *Id.*

that a policy provision is ambiguous or admitting *evidence* regarding the meaning of the provision. For example, in *Port Authority of New York and New Jersey v. Affiliated FM Ins. Co.*, the court held that the presence of asbestos fibers in buildings operated by the policyholder did not constitute “physical loss or damage” under its property insurance policies, because the buildings were not uninhabitable or unusable. Interpreting the provision to provide coverage, the court said, “would not comport with the intent of a first-party ‘all-risks’ policy, but would transform it into a maintenance contract.”⁷² There is nothing in the opinion indicating that evidence regarding the “intent of a first-party all risks policy” had been introduced, because there was no need for such evidence. That is a matter of legislative, not adjudicative fact. The court’s own knowledge of the function of property insurance, as distinguished from maintenance contracts, informed its interpretation of the policy provision and led it to the plain meaning of the policy.⁷³

The court in *City of Johnstown v Bankers Standard Ins. Co.*⁷⁴ relied on the purpose of the insurance policy at issue in that case in a very similar manner. The insured in that case claimed coverage of environmental cleanup liability under its CGL insurance policies. Its insurers denied coverage on the basis of a provision that excluded coverage of liability for property damage that was “expected or intended” by the insured, arguing that the provision precluded coverage of “risks” that the insured expected or intended.⁷⁵ The court rejected this argument, holding that “to exclude all losses or damages which might in some way have been *expected* by the insured could expand the field of exclusion until virtually no recovery could be had on insurance. This is so since it is mishaps that are ‘expected’ – taken in its broadest sense – that are insured against.”⁷⁶ This conclusion was obviously premised on the court’s understanding of the principal risks of liability that CGL insurance policies are intended to cover. In effect the court held that because CGL insurance policies are designed to cover liability for

⁷² 311 F.3d 226, 236 (3d Cir. 2002).

⁷³ *Id.* There may of course be cases in which the purpose of a particular policy provision is not a legislative fact, because the purpose arises out of a feature of industry custom, or the needs of a particular policyholder. But the general purpose of a particular form of insurance will almost always be a legislative, not an adjudicative, fact.

⁷⁴ 877 F.2d 1146 (2d Cir. 1989).

⁷⁵ *Id.* at 1149.

⁷⁶ *Id.* at 1150 (*italics in original*).

negligence, the “expected or intended” provision could not properly be interpreted to exclude coverage of liability for most negligence.

D. TAKING STOCK

Courts subscribing to the plain-meaning rule routinely consider certain matter outside the four corners of the insurance policy when interpreting policy provisions: implicit knowledge, adjudicative facts subject to judicial notice, and legislative facts. I think that the best way to understand this practice is not to consider each source an exception to the rule that extrinsic evidence is not admissible if a policy provision has a plain meaning. Rather, although these sources of meaning are extrinsic to the insurance policy, they are not extrinsic evidence because, for all practical purposes, they are not *evidence*. By this I mean not only that they need not be formally admitted into evidence. More importantly, what renders them not *evidence* is that their existence or non-existence is not a question of fact that is subject only to the highly deferential review that is accorded to findings of fact at the appellate level.

On the contrary, whether to employ implicit knowledge, take judicial notice of an adjudicative fact, or rely on a legislative fact – and what these sources reveal to be true – are decisions for the court – in effect, decisions of law – and as such are subject to *de novo* review on appeal. This insight explains and justifies the vast majority of references to and reliance on sources outside the four corners of the insurance policy by plain-meaning courts.

This is especially important in view of the fact that the vast majority of insurance disputes concern standard-form policy provisions whose meaning, whatever it is, governs the rights of numerous policyholders. If standard-form policy provisions are to have a standard meaning, the resolution of insurance disputes must have the effect of precedent under *stare decisis*. If interpretations typically involved the resolution of questions of fact, then most interpretive decisions by the courts that relied on judicial notice and legislative facts could have little or no *stare decisis* effect. But of course they do have that effect. It follows that judicial interpretations of standard-form policy provisions relying on sources of meaning extrinsic to the policy are not resolutions of questions of fact based on conventional evidence, and therefore that the sources of meaning outside the four corners of the policy on which they rely are not really factual evidence at all. In short, there is nothing inconsistent or paradoxical about plain-meaning courts considering these sources of meaning.

IV. AMBIGUITY ABOUT AMBIGUITY: INTERPRETATION AND
CONSTRUCTION OF AMBIGUOUS POLICY PROVISIONS

Once a plain-meaning court determines that a policy provision is ambiguous, then the plain-meaning rule no longer precludes the admission of extrinsic evidence to determine the more reasonable interpretation of the ambiguous provision.⁷⁷ Only the conventional rules of evidence limit what the court may consider when a policy provision is ambiguous. Common forms of extrinsic evidence include the negotiations, if any, between the parties, their course of dealing once the policy has been issued,⁷⁸ custom and usage,⁷⁹ and the drafting history of standard-form policy provisions.⁸⁰

In most jurisdictions the courts first attempt to determine whether, in light of any extrinsic evidence that is admitted after a court holds that a policy provision is ambiguous, the ambiguous provision has a single meaning.⁸¹ This is interpretation. Only if the extrinsic evidence does not resolve the ambiguity in this fashion does the court then apply *contra proferentem*. This is construction. An important but little-recognized issue regarding interpretation of ambiguous policy provisions is the *stare decisis* effect of an interpretation, for this concerns whether, and when,

⁷⁷ RLLI § 4 cmt. b.

⁷⁸ See, e.g., *Motors Liquidation Co. v. Allianz Ins. Co.*, No: N11C-12-022 FSS CCLD, 2013 WL 7095859 at *5 (Del. Super. Ct. Dec. 31, 2013).

⁷⁹ See, e.g., *Transp. Indem. Co. v Dahlen Transp., Inc.*, 161 N.W. 2d 546, 550 (Minn. 1968) (considering evidence of insurance industry custom and usage).

⁸⁰ See, e.g., *Balandran v. Safeco Ins. Co. of Am.*, 972 S.W.2d 738, 741-42 (Tex. 1998).

⁸¹ In a minority of jurisdictions, the courts apply *contra proferentem* immediately upon concluding that a policy provision is ambiguous, without considering extrinsic evidence of the provision's meaning. See, e.g., *Wash. Nat'l Ins. Corp v. Ruderman*, 117 So. 3d 943, 949 (Fl. 2013); Daniel Schwarcz, *Coverage Information in Insurance Law*, 101 MINN. L. REV. 1457, 1505-10 (arguing that this should be the rule generally). In addition, sometimes when a court immediately applies *contra proferentem*, it is not possible to determine which approach a court is applying, because it is unclear from the court's opinion whether any extrinsic evidence was introduced. See, e.g., *Vargas v. Ins. Co. of N. Am.*, 651 F.2d 838 (3d Cir. 1981); *N. Pac. Ins. Co. v. Hamilton*, 22 P.3d 739 (Or. 2001).

interpretations of unambiguous and ambiguous policy provisions have the same kind of precedential effect. Similarly, a little-recognized issue regarding the construction of ambiguous policy provisions is whether there is a role to be played by the very different reasons for a policy provision's ambiguity. I discuss both issues below.

A. INTERPRETING AMBIGUOUS POLICY PROVISIONS

Sometimes interpretation based in part on extrinsic evidence reaches the conclusion that the policy provision at issue has a single meaning. Interpretations of this sort typically are given the same strong *stare decisis* effect as interpretations under the plain-meaning rule. But whether this makes sense depends on what sources of meaning were called upon to interpret the provision. If the interpretation is a decision of law, then it should have that *stare decisis* effect. On the other hand, if truly evidentiary, factual sources have been considered – as is permitted once the provision has been determined to be ambiguous – then the interpretation arrived at may be based in whole or in part on findings of fact, and the *stare decisis* effect of the decision should be more limited.⁸²

For example, if in order to interpret a provision a court based its interpretation on statements made by the parties in negotiating the policy or in custom-drafting it, or on the course of dealing between the parties subsequent to the issuance of the policy, then the decision would have only the *stare decisis* effect, if any, that a decision relying on findings of fact about such matters may have. A dispute between different parties over the meaning of the same ambiguous policy provision could therefore be resolved differently, depending on the extrinsic evidence relevant to their independent dispute.⁸³ An appellate court reviewing such interpretations would have only

⁸² Professor Farnsworth has made this point about the scope of review of contract decisions generally. See FARNSWORTH, *supra* note 15, §7.17. But he does not extend the point to *stare decisis*, and I have never seen the point made, in either respect, about interpretations of insurance policy language.

⁸³ For examples of such disputes in classic cases, see *Silberg v. Cal. Life Ins. Co.*, 521 P.2d 1103, 1109 (Cal. 1974) (involving dispute over insurance industry custom regarding payment of claims and subsequently seeking reimbursement from insured or another insurer); *Morton Int'l, Inc. v. Gen. Acc. Ins. Co.*, 629 A.2d 831, 855 (N.J. 1993) (involving dispute over inferences to be drawn from drafting history of the qualified pollution exclusion in CGL policies).

limited authority to reverse them, presumably for something like abuse of discretion.

On the other hand, suppose that a court rested its interpretation on findings of fact that were more generally relevant to the interpretation of the policy provision, such as custom and usage within an industry (including the insurance industry), or the drafting history of a standard-form provision. There is no question that appellate courts treat such interpretations as being reviewable *de novo*.⁸⁴ In my experience, subsequent courts asked to interpret the same policy provision treat the earlier decision as having *stare decisis* effect. This evidently precludes revisiting the factual predicates on which the earlier decision rested, including the nature and significance of custom and usage, and implications of the drafting history of the ambiguous policy provision at issue.

Why do interpretations resting on findings about generally-applicable facts such as industry custom and usage, and standard-form drafting history, have the broad *stare decisis* effect that they are usually accorded, even though the interpretations rest on factual premises that in other settings could be relitigated? For example, why are interpretations based on drafting history treated as if they are not subject to re-litigation in a claim by a different policyholder?

I think there are several possible explanations. First, this treatment may actually be unjustified or not even (strictly speaking) what actually happens. In fact, it may be that courts adhere to decisions based on findings of fact regarding industry custom or drafting history mainly because subsequent litigants do not attempt to introduce new evidence regarding these matters, and that the courts would actually consider substantial new evidence, and decide differently if the evidence warranted doing so, especially after the passage of a considerable amount of time. Perhaps decisions based on such matters of fact are potentially subject to re-litigation.

Second, however, if this is the case, re-litigation could be highly disruptive. Then, a particular trial court's findings of fact about custom or drafting history might in principle be subject to a different inference by a subsequent trial court, even if no new evidence were introduced. Standard policy language would then potentially be subject to different interpretations, depending on findings by different courts, case-by-case.

⁸⁴ See, e.g., *PBM Nutritionals, LLC v. Lexington Ins. Co.*, 724 S.E.2d 707, 712 (Va. 2012) (holding that an appellate court reviews interpretations of insurance contracts *de novo*).

Third, the courts may be treating findings about such matters as custom and usage and drafting history as involving legislative rather than adjudicative facts, and therefore subject to *de novo* appellate review that generates a strong *stare decisis* effect. That may be an accurate characterization of certain such facts – the general and undisputable explanation for the addition of an absolute pollution exclusion to CGL insurance policies beginning in 1986 is a good example.⁸⁵ But that cannot be the explanation for the strong *stare decisis* effect accorded other interpretations that involve disputable characterizations of custom and usage or drafting history.

Fourth, the justification for this treatment may resonate with the principle underlying non-mutual offensive collateral estoppel. If a policyholder has won a case involving facts generally relevant to the meaning of a standard-form policy provision, then a subsequent, different insurer is treated as being bound by the earlier decision that technically binds only the earlier insurer. The second insurer stands in the shoes of the earlier insurer as long as the earlier insurer had an incentive to fully litigate the issue in question, and lost. This whole analogy would work against insurers, but not against policyholders, since one policyholder cannot reasonably be understood to have been litigating on behalf of all policyholders.

Finally, it may be that the practice of according *stare decisis* effect to interpretations based on generally-applicable facts is a prudential exercise by the courts rather than one that, technically, is mandatory. The practice facilitates treating standard-form policy language as having a uniform meaning, and thereby enhances the advantages of having standard-form policies.⁸⁶ If an earlier factual finding about the significance of custom and usage or drafting history for the meaning of an ambiguous standard-form provision turns out later to have been flawed, later courts still can exercise their discretion to revisit it. Since *stare decisis* is itself an essentially prudential doctrine,⁸⁷ the seemingly anomalous precedential effect that the courts give to interpretations based on custom and usage and drafting history may be more apparent than real.

⁸⁵ See *Am. States Ins. v. Koloms*, 687 N.E.2d 72, 79-82 (Ill. 1997).

⁸⁶ See RESTATEMENT (SECOND) OF CONTRACTS § 211(2) (indicating that interpreting standard form contracts so as to treat similarly-situated parties in the same manner is desirable).

⁸⁷ See FREDERICK SCHAUER, THINKING LIKE A LAWYER 36 (2009) (noting the practice of “ordinarily requiring that decisions follow precedent”).

B. CONTRA PROFERENTEM: THREE LEVELS OF CONSTRUCTION
AGAINST THE DRAFTER

It is quite possible for a policy provision to remain ambiguous even after consideration of extrinsic evidence relevant to the meaning of the provision. In such cases, under the traditional application of *contra proferentem*, the provision is construed against the drafter, which in the case of standard-form provisions is the insurer. The one recognized limit in this situation is that a construction that affords policyholders coverage that they could not reasonably expect is not to be adopted.⁸⁸ This might be understood as an interpretive limit on *contra proferentem*, since in a sense an interpretation that would afford policyholders more coverage than they would reasonably expect is not a reasonable interpretation.

There also are hints in the case law, however, that the reason that a policy provision is ambiguous may have a bearing on the process of construction. These hints actually have a substantial normative basis that has not been recognized: the greater the amount of blame for the ambiguity that can be attributed to the drafter, the stronger the justification for construing the provision against the insurer and in favor of coverage. There are three levels of blameworthiness, corresponding to the reason that the provision came to be drafted as it was.

1. Ambiguity by Necessity

Some policy provisions are ambiguous out of necessity. The problem they address may be complex, the language that would be required to unambiguously resolve particular issues in advance of all disputes may be lengthier than is practical or desirable,⁸⁹ or all the situations to which the

⁸⁸ See RLLI, at § 4 cmt. f; *Chute v. North River Ins. Co.*, 214 N.W. 473, 474 (Minn. 1927) (holding that construing an ambiguous policy provision in favor of coverage the policyholder could not reasonably expect would “ignore the purpose of the contract”). In my earlier Article about insurance policy interpretation, I called this the “majoritarian” approach, distinguishing it from a “penalty” approach that would construe an ambiguous provision in favor of coverage regardless of whether it was reasonable to expect that coverage. Abraham *supra* note 14, at 545-50.

⁸⁹ See Daniel Schwarcz, *Coverage Information in Insurance Law*, 101 MINN. L. REV. 1457, 1474 (2017) (arguing that longer policy provisions

provision may apply may be difficult to predict and unambiguously address. As one court astutely put this point:

Drafters cannot anticipate all possible interactions of fact and text, and if they could to attempt to cope with them in advance would leave behind a contract more like a federal procurement manual than like a traditional insurance policy. Insureds would not be made better off in the process. The resulting contract would not only be incomprehensible but also more expensive.⁹⁰

For example, the standard-form homeowners policy defines an insured (among other things) as “residents of your [the policyholder’s] household who are...²¹ and in your *care* or the *care* of a resident of your household who is your relative.”⁹¹ Although I have found no cases on the issue, there is a pretty good argument that the word “care” in this context is ambiguous, in the sense that it has either pretty broad or quite narrow boundaries.⁹² Being “in your care” might require that the policyholder serve as the complete support for a bedridden person who is unable to perform bodily functions without assistance, or it might require only serving a temporarily ill person meals and helping the person to get out of bed. But specifying the exact contours of “care,” especially given the different possible gradations of “care,” would require extended verbiage addressed to an issue that is likely to arise only rarely under homeowners’ policies. The result is that the term “care” is ambiguous, and would likely remain

makes it more difficult for policies to be understood *ex ante*, by those selling them, regulating them, and deciding whether to buy them). In my earlier article, I referred to this kind of assessment as the application of a “linguistic standard of care.” Abraham, *supra* note 14, at 537-38. To ignore this factor would, I think, be to impose strict liability on the drafter for employing unavoidably ambiguous policy language. *Id.* at 538-40.

⁹⁰ See *Harnischfeger Corp. v. Harbor Ins. Co.*, 927 F.2d 974, 976 (7th Cir. 1991); FARNSWORTH, *supra* note 15, §7.09 (indicating that the difficulty of foreseeing all the circumstances that will arise sometimes accounts for lack of clarity).

⁹¹ See ABRAHAM & SCHWARCZ, *supra* note 1, at 186-87 (emphasis added).

⁹² Professor Farnsworth identifies haziness at the boundary of a concept as “vagueness,” and suggests that intentional vagueness may be more justified than intentional ambiguity. FARNSWORTH, *supra* note 15, at §7.09.

ambiguous even if extrinsic evidence addressing the meaning of the term were available and admitted.

Similarly, the standard-form CGL policy covers liability incurred because of bodily injury or property damage that occurs “during the policy period,” and further provides that such bodily injury or property damage “includes any continuation, change or resumption of *that* ‘bodily injury’ or ‘property damage’ after the end of the policy period.”⁹³ The term “that” (technically, a “demonstrative adjective”⁹⁴) is almost certainly ambiguous in some contexts. Suppose that during policy year one, hazardous waste leaks from a site and contaminates groundwater (underground water) lying fifty feet beyond the boundary of the site where it was deposited. That is “property damage.” Suppose further, however, that in policy year two, the waste that was already in the groundwater migrates further, and contaminates previously-uncontaminated groundwater, lying between 50 and 500 feet beyond the boundary of the site. Under the above-quoted provision, is the contamination that occurred during policy year 2 a continuation of “that” original property damage (in which case it is not covered under the policy in force during year 2), or is it new “property damage” that is not counted as part of the original property damage (and therefore is covered under the policy in force in year 2)?

In this context, both are arguably reasonable interpretations on the face of this policy language. Extrinsic evidence would be admissible to determine which interpretation is more reasonable, but may well not resolve the ambiguity. Yet, whichever interpretation is adopted would have required extensive verbiage to address unambiguously in the policy, especially since my hypothetical is an example of only a number of different factual scenarios that might have to be identified. In effect, the term “that” is a placeholder that, understandably, delegates the task of elaboration to the courts.

This necessity explanation for ambiguity of the sort reflected in the homeowners’ policy’s use of the term “care,” and the CGL policy’s use of the term “that,” justifies an evenhanded search for the more reasonable interpretation of the policy provision in the context of the claim at issue, because the insurer’s drafting does not reflect sloppiness or an effort to take advantage of policyholders. In this setting, *contra proferentem* should

⁹³ *Id.* at 439 (emphasis added).

⁹⁴ See *What is a Demonstrative Adjective?*, YOUR DICTIONARY <https://grammar.yourdictionary.com/parts-of-speech/adjectives/demonstrative-adjectives.html> (last visited on May 20, 2019).

operate as a rule of last resort, a genuine tiebreaker to be used only when the evidence does not generate a single interpretation that is at least slightly more reasonable than a competing interpretation or interpretation.

The courts, however, tend not to resolve such situations in this way. Rather, often they avoid holding that a policy provision that is ambiguous out of necessity, and instead treat the situation as calling for the creation of a rule governing the problem rather than for an interpretation of the policy language. Perhaps the most prominent example of this approach is the courts' adoption of the pro-rata approach to the allocation of coverage responsibility among multiple triggered CGL insurance policies that were issued before the above-quoted provision addressing the continuation of injury or damage was included in the standard-form policy. The CGL policy covered liability "for those sums that the insured becomes legally obligated to pay as damages... because of bodily injury or property damage...which occurs during the policy period."⁹⁵ The policy did not unambiguously address the extent of each policy's coverage responsibility if bodily injury or property damage occurred during multiple policy periods. But the courts nonetheless developed a rule that allocated coverage responsibility on a pro-rata basis.⁹⁶

Another example of this approach is the courts' treatment of exclusions from coverage of liability for bodily injury or property damage "expected or intended from the standpoint of the insured."⁹⁷ These exclusions are arguably ambiguous in a number of ways. For instance, they do not address whether coverage of one insured is excluded when another insured expected or intended harm; and they do not address whether coverage is excluded when one type of harm (*e.g.*, bodily injury) is expected but a different type of harm (*e.g.*, property damage) occurs. On one view, addressing all these possibilities would render the provision unduly complicated.⁹⁸ Instead of holding that the provision is ambiguous, therefore,

⁹⁵ See KENNETH S. ABRAHAM, *INSURANCE LAW & REGULATION* 440 (1990).

⁹⁶ See, *e.g.*, *Boston Gas Co. v. Century Indem. Co.*, 910 NE. 2d 290 (Mass. 2009). Other courts adopted a rule of joint and several responsibility. See, *e.g.*, *Goodyear Tire & Rubber Co. v. Aetna Cas. & Sur. Co.*, 769 N.E. 2d 835 (Ohio 2002).

⁹⁷ ABRAHAM & SCHWARCZ, *supra* note 1, at 440.

⁹⁸ The reason I have qualified this statement is that the standard-form homeowners policy has addressed this issue with fairly straightforward language. See *id.* at 205 (setting out a provision indicating that liability for harm is excluded even if that harm is "of a different kind, quality or degree

the courts have developed rules (though they vary) addressing these permutations.⁹⁹

In such cases, the courts do not appear to be engaging in either interpretation or construction. In fact, however, they are doing the latter. The best understanding of the courts' development of rules of this sort to see it as the construction of policy provisions that are necessarily ambiguous. Because the drafter of provisions that are ambiguous by necessity is arguably not to blame for such ambiguity, invoking *contra proferentem* is regarded as inappropriate. Instead, the courts substitute a rule for what would otherwise be an unduly complex policy provision.

2. Ambiguity Resulting from Faulty Drafting

A second, and more blameworthy, reason that a policy provision may ambiguous is that it has been poorly drafted. The courts have frequently held that, the more easily it would have been to draft a provision that would have rendered its meaning clear, the stronger the argument that the provision is ambiguous.¹⁰⁰ The classic statement of this notion is that of Judge Frankel in the *Pan American* case:

Where the risk is well known and there are terms reasonably apt and precise to describe it, the use of substantially less certain phraseology upon which dictionaries and common understanding may fairly differ, is likely to result in interpretations favoring coverage rather than exclusion.¹⁰¹

than initially expected for intended").

⁹⁹ See, e.g., *SL Indus., Inc. v. American Motorists Ins. Co.*, 607 A.2d 1266 (N.J. 1992) (preserving the possibility of coverage when a different type of harm than occurred was expected); *American Family Ins. Co. v. Walser*, 628 N.W. 2d 605 (Minn. 2001) (holding that the exclusion applies as long as some harm was expected).

¹⁰⁰ See *ABRAHAM & SCHWARCZ*, *supra* note 1, at 540-44 (referring to this as a "perfectibility standard").

¹⁰¹ *Pan American World Airways, Inc. v. Aetna Cas. & Sur. Co.*, 368 F. Supp. 1098, 1188 (S.D. N.Y. 1973), *aff'd*, 505 F.2d 989 (2d Cir. 1974). See also *Estrin Const. Co. v. Aetna Cas. & Sur. Co.*, 612 S.W. 2d 413 (noting that an inept drafter has the resources to do better).

For example, in *Vlastos v. Sumitomo Marine & Fire Ins. Co. (Europe)*, a policy provisions stated, “Warranted that the 3rd floor is occupied as [a] janitor’s residence.”¹⁰² There was evidence that a janitor lived on the third floor, but that it had other uses as well. The insurer denied coverage on the ground that the provision required that the third floor be used *only* as a janitor’s residence. The court rejected this argument, holding that the provision was ambiguous its face, because “occupied” could reasonably have meant either “occupied exclusively” or “occupied in part.”¹⁰³ The fact that the addition of a single word – “exclusively” – would have rendered the provision unambiguous was central to the court’s reasoning.¹⁰⁴

Because one of the strongest arguments for *contra proferentem* has always been that the doctrine gives insurers the incentive to draft unambiguous policy provisions, a strong version of *contra proferentem* tends to be applied to faulty drafting that results in ambiguity. If an insurer drafts a sloppy or imprecise provision that could have been made unambiguous with little additional effort and no corresponding disadvantage, then the insurer has, in effect, been negligent. If there is also no extrinsic evidence supporting the insurer’s interpretation, the principle underlying *contra proferentem* strongly supports construing the provision against the drafter.

Even if there is extrinsic evidence supporting the insurer, however, *contra proferentem* should have substantial gravitational pull. The argument for rescuing the insurer by heavily weighing sources of extrinsic evidence such as the negotiations between the parties or custom and usage, is weak in this setting. Only strong and highly persuasive evidence should be permitted to rescue the insurer from its own faulty drafting in such a situation.

3. Ambiguity as a Trap

The last reason for ambiguity resonates most strongly with the policy underlying *contra proferentem*: the use of ambiguous policy language as a trap. Ordinarily there will be no direct evidence that an insurer deliberately

¹⁰² *Vlastos v. Sumitomo Marine & Fire Ins. Co. (Europe)*, 707 F.2d 775 (3d Cir. 1983).

¹⁰³ *Id* at 780.

¹⁰⁴ *See also* *Great American Fidelity Ins. Co. v. JWR Construction Services*, 882 F. Supp. 2d 1340, 1356 (S.D. Fl. 2012) (indicating that had the insurer “wished to exclude the faulty work of persons acting on [the insured’s] behalf, it could easily have done so by using clear policy language to that effect”).

designed an ambiguous policy provision with the aim of using it as a trap. Certainly standard-form language, drafted by committees of organizations such as ISO, rarely if ever has that aim. But individual insurers sometimes draft language that sets a trap, and even standard-form language is sometimes seized upon by individual insurers in a manner that functions like a trap.

For example, in *Vargas v. Ins. Co. of N. Amer.*, an aviation insurance policy covered occurrences, accidents, or losses that happened “within the United States of America, its territories or possessions, Canada or Mexico.”¹⁰⁵ The insured’s plane crashed in the sea, twenty-five miles west of Puerto Rico, on a flight that began in New York, with stops in Miami and Haiti.¹⁰⁶ The insurer denied coverage, on the ground that the loss did not occur “within” the required territory. Yet the insurer knew that the insured planned to fly the plane in the Caribbean. This was a blatant effort to use the literal meaning of the word “within” to avoid coverage. The court rejected that effort, holding that the word “within” was ambiguous – subject to more than one reasonable interpretation -- and construed the provision against the insurer, in part because the insured’s interpretation was consistent with the “realities of airplane travel,” which sometimes requires flights between two places within the continental United States (such as flights between New York and Miami) to “pass over waters beyond the territorial limits” of the United States.

The insurer’s attempt to set a trap was even more blatant in *Silberg v. California Life Ins. Co.*,¹⁰⁷ the seminal decision permitting the imposition of extracontractual liability on an insurer for bad-faith denial of a claim. There a provision in a health insurance policy excluded coverage of “any loss caused by or resulting from (1) injury or sickness for which compensation is payable under any Workmen’s Compensation...Law.”¹⁰⁸ The insurer asserted that this exclusion precluded coverage, not only of any loss paid by workers’ compensation but of all losses incurred by the insured for an injury or sickness for which workers’ compensation paid anything at all.

The insurer therefore refused to pay any of the insured’s losses until it was determined whether any workers’ compensation would be paid, and when the insured settled his workers’ compensation claim, the insurer denied

¹⁰⁵ See *Vargas v. Ins. Co. of N. Am.*, 651 F.2d at 840.

¹⁰⁶ *Id.*

¹⁰⁷ 521 P. 2d 1103 (Cal. 1974) (emphasis added).

¹⁰⁸ *Id.* at 1111.

coverage entirely, including for the losses that workers' compensation did not pay. The insurer did this despite having advertised "ALL BENEFITS PAYABLE IN FULL REGARDLESS OF ANY OTHER INSURANCE YOU MAY HAVE."¹⁰⁹ This is about as clear an example as there can be of insurer's attempt at bait-and-switch. Clearly the insurer was using the ambiguity of the word "payable" as a trap.

Finally, the insurer in *Corban v. United Automobile Services Association*¹¹⁰ took the position that the "anti-concurrent causation" clause in its homeowners policy, which precluded coverage if a loss was "caused" by an excluded cause, even if a covered cause contributed "in any sequence" to the loss, precluded coverage of loss caused by hurricane wind damage, if later-caused (and excluded) water damage contributed to that loss.¹¹¹ The court held that the phrase "in any sequence" was ambiguous and invoked *contra proferentem*, in part because the insurer's interpretation would have excluded coverage of loss that had already occurred, on the ground that an excluded cause subsequently contributed to the loss.¹¹² The insurer's position offered what in common parlance would be called a "gotcha" interpretation: "your loss occurred, and it was covered at the moment it occurred, but subsequent events out of your control deprived you of insurance for this already-covered loss." Few if any insureds would have expected that result, or understood the relevant language to provide for it. It reflected an effort by the insurer to use the language of the anti-concurrent causation clause as a trap that had been set for the insured and then sprung after a covered loss had occurred.

CONCLUSION

The interpretation of insurance policies turns out to be a more sophisticated and more complex process than the myths about interpretation sometimes make it out to be. Interpretations yielding the conclusion that a policy provision has a single, plain meaning are often active searches for meaning, not passive receptions of a meaning that is evident on the face of a provision. In addition, the idea that extrinsic evidence is not admissible to aid in the interpretation of a provision with a plain meaning masks the fact that important sources of meaning that are outside the policy are commonly considered in the course of interpretation. Finally, both the *stare decisis*

¹⁰⁹ *Id.*

¹¹⁰ 20 So. 3d 601 (Miss. 601).

¹¹¹ *Id.* at 612.

¹¹² *Id.* at 615.

effect of interpretations of ambiguous policy language, and differences in the reasons that an ambiguous policy provision came to be included in a policy, pose issues that have gone largely unrecognized because of the tendency to oversimplify what occurs when ambiguous policy language is interpreted or construed. Greater clarity about all of these interpretive phenomena deepens our understanding of the role played by plain meaning, extrinsic evidence, and ambiguity in the interpretation of insurance policies.

