

UNITED STATES COURT OF APPEALS

FOR THE SECOND CIRCUIT

August Term, 2003

(Argued: March 17, 2004

Decided: August 22, 2005)

Docket No. 03-1356

UNITED STATES OF AMERICA,

Appellee,

—v.—

MANUEL GONZALEZ, also known as “Luis Enrique Dirocie Bello,”¹
also known as “Tito,”

Defendant-Appellant,

ROBINSON JIMENEZ,

Defendant.

Before:

SACK, SOTOMAYOR, and RAGGI, *Circuit Judges.*

¹ At a hearing before the district court, the defendant testified that his real name is “Luis Enrique Dirocie Bello.” Because the judgment is entered only in the name “Manuel Gonzalez,” however, and because the parties continue to refer to the defendant by that name in their submissions to this court, we do likewise in this opinion. Nevertheless, we have directed the Clerk of the Court to amend the caption of this case to indicate that Gonzalez is also known by the name Dirocie Bello.

Appeal from a judgment of conviction entered in the United States District Court for the Southern District of New York (John S. Martin, Jr., *Judge*), which sentenced defendant to a mandatory minimum term of twenty years' incarceration for conspiracy to distribute fifty grams or more of crack cocaine, on the ground that defendant's guilty plea was inadequate to support conviction on an aggravated drug charge because he had been misinformed as to his right to have the statutory drug quantity proved to a jury rather than the court and he had not admitted that quantity in his allocution.

REMANDED with instructions to vacate the judgment and allow the defendant to withdraw his guilty plea.

TRACY W. YOUNG (Lisa Scolari, *on the brief*), New York, New York, *for Defendant-Appellant*.

ROBERTO FINZI, Assistant United States Attorney (Neil M. Barofsky, Assistant United States Attorney, *on the brief*), *for* David N. Kelley, United States Attorney for the Southern District of New York, New York, New York, *for Appellee*.

REENA RAGGI, *Circuit Judge*:

Defendant-appellant Manuel Gonzalez appeals from a judgment of conviction entered on June 4, 2003, in the United States District Court for the Southern District of New York (John S. Martin, Jr., *Judge*) based on defendant's guilty plea to a single-count indictment charging a conspiracy to distribute and possess with intent to distribute fifty grams or more of cocaine base, commonly referred to as "crack." See 21 U.S.C. §§ 812, 841(a)(1),

841(b)(1)(A), 846, 851. Gonzalez is presently serving a twenty-year term of incarceration, a downward departure from his 262-to-327 month Sentencing Guidelines range. That incarceratory term was not selected by the district court in the simple exercise of its departure discretion. Rather, the court concluded that its discretion was curbed by the minimum twenty-year prison term mandated by § 841(b)(1)(A) in light of two facts: the quantity of drugs found by the court based on a preponderance of the evidence and Gonzalez's prior New York State felony drug conviction.

On this appeal, Gonzalez does not raise a direct challenge to his sentence. Instead, he asserts that the district court erred in denying his pre-sentence motion to withdraw his guilty plea. Gonzalez argues that Apprendi v. New Jersey, 530 U.S. 466 (2000), and United States v. Thomas, 274 F.3d 655 (2d Cir. 2001) (en banc), both decided in the interim between his plea and sentence, obliged the district court to conclude that his guilty plea was not knowing, voluntary, or sufficient to support the crime of conviction because (1) he had not admitted the statutory drug quantity, and (2) he had been misinformed as to his right to have a jury rather than the court determine that quantity.

The government submits that Gonzalez's plea challenge lacks merit because quantity is not an element of a § 841(b)(1)(A) drug offense when, as in this case, a defendant receives a mandatory minimum sentence not in excess of the prescribed maximum for an identical unquantified crime pursuant to § 841(b)(1)(C). Both the structure of § 841 and our prior construction of that statute compel us to reject the government's argument.

Even if the right to trial, as recognized in Apprendi, is violated only by certain

sentences, the law cannot reasonably defer identification of the elements of a crime until after a prosecution is concluded. Thus, this court’s construction of the quantity provisions of § 841 as “elements” that have to be pleaded and proved to a jury or admitted by the defendant to support conviction on an aggravated drug offense under that statute,² see United States v. Thomas, 274 F.3d at 660, accord United States v. McLean, 287 F.3d 127, 134 (2d Cir. 2002); United States v. Outen, 286 F.3d 622, 635-36 (2d Cir. 2002); see also Coleman v. United States, 329 F.3d 77, 85-86 (2d Cir. 2003), while undoubtedly prompted by Apprendi, is not limited to post hoc review of particular sentences. We here clarify that the statutory drug quantity is an element in all prosecutions of aggravated § 841 offenses.

Although we dedicate considerable space in this opinion to explaining this point, its import on this appeal pertains directly to defendant’s guilty plea and only secondarily to his sentence. We conclude that Gonzalez’s guilty plea could not support conviction on a § 841(b)(1)(A) conspiracy without an admission to the drug quantity element of such an aggravated offense. Gonzalez made no such admission and, in fact, disputed the statutory quantity. Thus his plea at best supports conviction on a lesser, unquantified drug charge, whose sentencing range is prescribed by § 841(b)(1)(C). Because that sentencing range and

² In this opinion, we use the term “aggravated drug offense” to refer to crimes defined by reference to the lettered subsections of § 841(b)(1) that provide for enhanced penalties for drug trafficking in specified quantities. See 21 U.S.C. § 841(b)(1)(A)-(D). In Part II.B.1.a of this opinion, infra, we discuss in some detail how three subsections of § 841(b)(1) operate to proscribe crack trafficking generally without regard to quantity, see id. § 841(b)(1)(C), as well as to identify aggravated offenses for trafficking in specific quantities of crack, see id. § 841(b)(1)(A)-(B).

the higher ones provided in § 841(b)(1)(A) and -(b)(1)(B) for quantified crack offenses operate as unified and independent wholes, they cannot be deconstructed, as the government here suggests, so that drug quantity operates as an element for purposes of determining an applicable sentencing maximum but as a sentencing factor for purposes of determining the applicable mandatory minimum.

In sum, because drug quantity is an element of the aggravated offense of conviction, we conclude that Gonzalez was misinformed as to his right to have the statutory quantity proved to a jury. For that reason, and because his own allocution failed to provide an adequate factual basis on that element of an aggravated offense, his guilty plea could not be deemed knowing, voluntary, or sufficient to support a judgment of conviction on a § 841(b)(1)(A) charge. The government having refused to accept any lesser disposition of the case – as was its right – Gonzalez’s motion to withdraw his plea to a § 841(b)(1)(A) conspiracy should have been granted. We remand this case to the district court with directions that it vacate Gonzalez’s conviction, allow him to withdraw his guilty plea, and permit the government to proceed with the prosecution of the charged § 841(b)(1)(A) conspiracy. In doing so, the government may still agree to accept a lesser disposition with Gonzalez pleading guilty to an unquantified drug conspiracy under § 841(b)(1)(C), but, in that case, Gonzalez would be sentenced pursuant to the zero-to-thirty year range of imprisonment prescribed by that section, without regard to any § 841(b)(1)(A) mandatory minimum. See generally United States v. Yu, 285 F.3d 192, 198 (2d Cir. 2002).

I. Background

A. The Crack Conspiracy

The conspiracy at issue in this case involved an attempted drug sale on July 16, 1998. On that date, agents of the Drug Enforcement Administration (“DEA”) recorded a telephone conversation between one of their confidential informants and Gonzalez in which Gonzalez agreed to sell the informant one kilogram of crack cocaine for \$18,000. At a preliminary meeting, Gonzalez provided the informant with a sample of the crack to demonstrate its quality. Subsequent laboratory analysis confirmed the sample to be 0.4 grams of 87% pure cocaine base. Approximately one half-hour after the sample transfer, Gonzalez and the informant participated in another recorded telephone conversation and agreed to consummate their kilogram transaction later that same evening.

Around 8:30 p.m. on July 16, 1998, Gonzalez arrived at the appointed Manhattan meeting site in a car driven by co-defendant Robinson Jimenez. As soon as Gonzalez showed the informant a small bag containing what appeared to be a quantity of crack, the informant gave a pre-arranged signal to DEA surveillance agents. When the agents moved forward, however, Gonzalez and Jimenez fled the scene, thereby avoiding both arrest and seizure of the proffered drugs.

The two men were, in fact, arrested on August 11, 1998, at which time Jimenez made inculpatory oral and written statements acknowledging his own awareness that the July 16, 1998 deal had been for a kilogram of crack cocaine. On August 25, 1998, a federal grand

jury named Gonzalez and Jimenez in a single-count indictment charging a conspiracy to distribute and possess with intent to distribute fifty grams or more of cocaine base, in violation of 21 U.S.C. §§ 812, 841(a)(1), 841(b)(1)(A), and 846.³

B. Gonzalez's Guilty Plea

On April 14, 1999, Gonzalez pleaded guilty to the indictment, while specifically disputing the quantity of drugs involved in the charged conspiracy. Earlier that same day, the government had filed a prior felony information pursuant to 21 U.S.C. § 851 based on Gonzalez's 1997 New York State conviction for the attempted sale of four ounces of cocaine to an undercover officer. The effect of the filing was significant. If Gonzalez were convicted and sentenced for conspiring to traffic in fifty grams or more of crack pursuant to 21 U.S.C. § 841(b)(1)(A), as charged in the indictment, the prior felony information would raise his statutory sentencing range from a minimum ten-years-to-life imprisonment to a minimum twenty-years-to-life imprisonment. Before the district court, Gonzalez's counsel complained that the government's purpose in filing the prior felony information was to retaliate against her client for rejecting a plea agreement that would have required him to stipulate to a drug quantity that, under the Sentencing Guidelines, would expose him to approximately twelve and a half years' incarceration. Counsel emphasized that her client wished to challenge the charged drug quantity because of the "tremendous" impact that factor

³ Jimenez subsequently pleaded guilty to a two-count superseding indictment charging a conspiracy and an attempt to distribute crack. Sentenced to 120 months' imprisonment, he has not filed an appeal.

could have on the likely term of incarceration. Plea Tr. at 2. Thus, Gonzalez proposed to plead guilty to conspiring to distribute only the 0.4-gram crack sample, a quantity that, under then-applicable law, would keep him within the parameters of § 841(b)(1)(C) and expose him to a Guidelines sentence of approximately two years. Compare U.S.S.G. § 2D1.1(c) (Level 36 for at least 500 grams but less than 1.5 kilograms of cocaine base) with id. (Level 14 for at least 250 but less than 500 milligrams of cocaine base). He specifically declined to plead guilty to conspiring to distribute the charged fifty grams or more of crack, explaining, through counsel, that he had never intended to sell the informant a kilogram of real crack. Instead, his intent was to defraud his customer by having him pay for real crack while supplying him with a counterfeit substance.

Prior to beginning the plea allocution, the district court, relying on our pre-Thomas precedent, asked Gonzalez directly if he understood that, if the drug quantity issue was resolved against him by the court, he would face “a minimum of 20 years in jail.” Plea Tr. at 12. Gonzalez confirmed that understanding and indicated that he would “plead guilty [only to] giving a sample to the informant.” Id. at 13.

The district court proceeded to place Gonzalez under oath and to engage in a colloquy to ensure his competency, his awareness of his rights, and the voluntariness of his actions. In addressing the potential sentence resulting from a guilty plea, the district court relied on § 841(b)(1)(A), advising Gonzalez that “the maximum penalty, provided for in this statute is life imprisonment,” id. at 16; “lifetime supervised release,” id. at 17; a “\$4 million” fine, and a mandatory “special assessment of \$100,” id. at 18. The court reiterated that, under the

government's view of the drug quantity at issue, the court would be required to impose a mandatory minimum prison term of twenty years, and a mandatory minimum supervised release term of five years. Gonzalez confirmed his understanding of these sentencing consequences.

In response to questions posed by the court as to the factual basis for his plea, Gonzalez admitted reaching an agreement with another person to violate the narcotics laws and, pursuant thereto, delivering a sample of crack cocaine to the informant. The court accepted the plea and set a schedule for the parties to file papers on the disputed issue of drug quantity.

C. The District Court's Determination of Drug Quantity

On November 15, 2001, the district court held a sentencing hearing pursuant to United States v. Fatico, 579 F.2d 707 (2d Cir. 1978), to resolve the disputed issue of drug quantity. To support its contention that the charged conspiracy involved fifty grams or more of crack cocaine, the government relied, in part, on the recorded telephone conversations between Gonzalez and the informant discussing a one-kilogram transaction, agent testimony that Gonzalez had sold the informant real crack on multiple occasions over the course of a year, and Jimenez's inculpatory post-arrest statement.

To refute this evidence and support his own claim that the charged conspiracy never intended to distribute more than the 0.4-gram crack sample, Gonzalez took the stand and testified that, beginning in 1997, he routinely sold crack buyers a counterfeit, non-controlled

substance commonly referred to as “nacona” (a Spanish contraction for “de nada con nada,” meaning “[s]omething that doesn’t contain anything”). Hearing Tr., Nov. 15, 2001, at 10. Gonzalez stated that he had been arrested several times for crack dealing that, in fact, involved nacona.⁴ He insisted that on July 16, 1998, he provided the DEA informant with a small sample of crack only to induce him to pay for a kilogram of nacona that would masquerade as genuine crack.

At the conclusion of the hearing, Gonzalez’s counsel argued that the government had failed to carry its burden of proof to establish a conspiracy to traffic in one kilogram of real crack cocaine. The government argued, however, that Gonzalez’s account of a counterfeit crack scheme was implausible because he could never have established a credible reputation as a drug dealer if he had repeatedly sold counterfeit drugs.

The able trial judge, whose considerable experience assessing criminal evidence predated his appointment to the federal bench, characterized this as “a very difficult case,” resolution of which could well turn on “the burden of proof.” *Id.* at 66. The judge specifically noted that, if the question of drug quantity had been presented to him “in a

⁴ Gonzalez’s testimony and defense exhibits showed that, in January 1997, he was arrested for third-degree sale of a controlled substance, only to have the charges against him dismissed when the police discovered that the substance sold was not, in fact, crack cocaine. Subsequently, in July 1997, he was charged with second-degree robbery after a disgruntled customer to whom he had sold nacona complained to the police. These charges were dropped when the complaining witness failed to appear to testify. Finally, in August 1997, Gonzalez was charged with third-degree sale of a controlled substance for what turned out to be a one-gram sample of crack cocaine and two kilograms of nacona. The evidence apparently did not reveal the disposition of this last case.

nonjury trial, with the burden of proof being beyond a reasonable doubt,” it would be difficult to conclude that the conspiracy involved a kilogram of crack. Id. After taking the matter under advisement, the district judge, on March 7, 2002, reported that he “did not find Mr. Gonzalez to be a credible witness,” a conclusion supported by reference to particular inconsistencies and implausibilities in the record. Hearing Tr., Mar. 7, 2002, at 2. He stated that, “on the preponderance of the evidence,” he was satisfied “that these defendants were conspiring to sell [one kilogram of] real narcotics.” Id.

D. Gonzalez’s Motion to Withdraw His Guilty Plea

Gonzalez subsequently moved to withdraw his guilty plea, arguing that his allocution did not knowingly or sufficiently support a conviction under 21 U.S.C. § 841(b)(1)(A) because he had not admitted the drug quantity necessary to trigger that statute’s enhanced penalty provisions and he had been misinformed as to his right to have that issue proved beyond a reasonable doubt to a jury rather than by a preponderance of the evidence to the court. At a January 23, 2003 status conference, the district court advised the parties that it was “considering granting the application.” Hearing Tr., Jan. 23, 2003, at 2. The court explained:

I am troubled about the allocution, whether I should not have made it more clear to the defendant his specific offense to which he was pleading. [Defense counsel], prior to my taking the plea, did make reference to 841(b)(1)(C). . . . [But] I think I should have probably gone into more detail with the defendant himself concerning the specific violations and what he was exposing himself to.

. . . .

I mean he certainly allocuted to the offense. However, he allocuted to two

offenses, because it's both an [841(b)(1)] A and [an 841(b)(1)] C violation. That should have been made clear. What I'm saying is I should have made clear to him that he was by that plea admitting a violation which if I found otherwise, that the quantity [of crack] was greater than [0.4 grams], he would be subjected to this more Draconian sentence.

Id. at 2-3, 4-5.

Upon receipt and review of further submissions from the parties, however, the district court, on April 17, 2003, denied Gonzalez's motion to withdraw his guilty plea. Rejecting defense counsel's argument that Gonzalez's failure to allocute to the drug quantity specified in § 841(b)(1)(A) precluded application of that statute's mandatory minimum sentence, the district court stated:

I don't think that is the law today, and that being the case, I really find no basis on which to set aside the plea. And, certainly, this defendant was aware that he was going to face this 20-year mandatory minimum sentence if I concluded that he was aware what he was selling in fact was crack and not simply Tylenol. So despite my original inclination, upon reflection, I don't find a basis on which I can legitimately say that there is a defect in the plea allocation, and certainly given the fact that the defendant litigated this all the way through a Fatico hearing, I don't think that there is any basis to say that it is appropriate, therefore, to allow him to withdraw his plea He has clearly admitted that he is guilty of these offenses, and I see no interest in justice allowing him to withdraw his plea at this time.

Hearing Tr., Apr. 17, 2003, at 3.

E. Sentencing

The district court sentenced Gonzalez on May 30, 2003. Based on its drug quantity finding, as well as its conclusion that Gonzalez had attempted to obstruct justice at the Fatico hearing, the court calculated his total base offense level at 38, which, with a criminal history

category of II, yielded a sentencing range of 262 to 327 months' incarceration. The court decided that a downward departure from this range was warranted "because . . . [its] finding [of drug quantity] was not beyond a reasonable doubt and while it satisfied the preponderance of the evidence standard, . . . given the magnitude of the adjustment resulting from that determination, it [was] appropriate to depart from the guideline." Sentencing Tr., May 30, 2003, at 2.⁵ Nevertheless, the court concluded that its departure discretion was necessarily curtailed by § 841(b)(1)(A), whose "mandatory minimum trumps the guideline at 240 months." Id. Finding that "a reduction at least to that level would be appropriate," id., the court sentenced Gonzalez to 240 months' imprisonment, ten years' supervised release, and a \$100 special assessment.

Gonzalez timely appealed his conviction.

II. Discussion

A. Standard of Review

Federal Rule of Criminal Procedure 11(d) states that "[a] defendant may withdraw a plea of guilty . . . (2) after the court accepts the plea, but before it imposes sentence if . . . (B) the defendant can show a fair and just reason for requesting the withdrawal."⁶ "The decision

⁵ The parties do not dispute the district court's Guidelines calculation, nor has the government cross-appealed its departure decision.

⁶ Rule 11 has been revised twice since Gonzalez entered his plea. The revisions did not effect any substantive change relevant here to provisions governing entering a plea. The revisions added provisions governing withdrawing a plea, formerly located at Rule 32, but

to allow a guilty plea to be withdrawn is committed to the discretion of the district judge” United States v. Torres, 129 F.3d 710, 715 (2d Cir. 1997) (describing predecessor rule, Fed. R. Crim. P. 32(e)); accord United States v. Harrington, 354 F.3d 178, 183 (2d Cir. 2004). We review the exercise of that discretion deferentially and will not reverse unless the denial of a withdrawal motion rests on an error of law, a clearly erroneous finding of fact, or otherwise “cannot be located within the range of permissible decisions.” Parker v. Time Warner Entm’t Co., L.P., 331 F.3d 13, 18 (2d Cir. 2003) (quoting Zervos v. Verizon N.Y., Inc., 252 F.3d 163, 169 (2d Cir. 2001)) (defining abuse of discretion). We apply the same standard of review to a district court’s finding that a defendant’s factual admissions support conviction on the charge to which he is pleading guilty. See United States v. Smith, 160 F.3d 117, 122 (2d Cir. 1998).

For reasons discussed herein, we are obliged to conclude that the denial of Gonzalez’s motion to withdraw his guilty plea rests on a legal error with respect to the proof necessary to establish the drug quantity element of an aggravated § 841 offense.

B. A Defendant Cannot Be Convicted of, and Is Not Subject to a Mandatory Minimum Sentence Pursuant to, § 841(b)(1)(A) or -(b)(1)(B) Unless the Statute’s Prescribed Drug Quantity Is Proved Beyond a Reasonable Doubt to a Jury or Admitted by the Defendant

_____ 1. Drug Quantity as an Element of an Aggravated Drug Offense

a. The Statutory Scheme

Title 21 U.S.C. § 841(a) makes it “unlawful for any person knowingly or intentionally

_____ this change took effect before Gonzalez moved to withdraw his plea. For ease of reference, we refer to Rule 11 as currently formulated.

– (1) to manufacture, distribute, or dispense, or possess with intent to manufacture, distribute, or dispense, a controlled substance,” such as crack cocaine, see 21 U.S.C. § 812 (identifying coca leaf derivatives as controlled substances pursuant to Schedule II(a)(4)). Any person who engages in conduct proscribed by § 841(a)(1), or who conspires to do so, see id. § 846, is subject to penalties detailed in the lettered subsections of § 841(b)(1). Three subsections pertain to persons who traffic in crack cocaine.

In cases involving “50 grams or more” of crack, § 841(b)(1)(A) prescribes the following incarceratory sentencing ranges:

- (1) a mandatory ten-year minimum to lifetime maximum for any offense in the prescribed quantity;
- (2) a mandatory twenty-year minimum to lifetime maximum if the drugs cause death or serious bodily injury or if the defendant has a prior final felony drug conviction;
- (3) a mandatory life sentence if the drugs cause death or serious bodily injury and the defendant has a prior final felony drug conviction; and
- (4) a mandatory life sentence without release if a defendant has two prior final felony drug convictions.

In cases involving “5 grams or more” of crack, § 841(b)(1)(B) prescribes the following ranges:

- (1) a mandatory five-year minimum to forty-year maximum for any offense in the prescribed quantity;
- (2) a mandatory ten-year minimum to lifetime maximum if the defendant has a prior final felony drug conviction;
- (3) a mandatory twenty-year minimum to lifetime maximum if the drugs cause death or serious bodily injury; and

- (4) a mandatory life sentence if the drugs cause death or serious bodily injury and the defendant has a prior final felony drug conviction.

In cases involving lesser or unquantified amounts of crack, § 841(b)(1)(C) prescribes the following ranges:

- (1) a zero minimum to twenty-year maximum for any offense;
- (2) a zero minimum to thirty-year maximum if the defendant has a prior final felony drug conviction;
- (3) a mandatory twenty-year minimum to lifetime maximum if the drugs cause death or serious bodily injury; and
- (4) a mandatory life sentence if the drugs cause death or serious bodily injury and the defendant has a prior final felony drug conviction.

As this brief review demonstrates, the statute is structured so that each lettered subsection operates independently of the others, without any need for cross-referencing. Within each subsection, the statute provides for each maximum sentence to be linked to a corresponding minimum (except where only a lifetime sentence is mandated). The independent operation of each sentencing range within each subsection is confirmed by the fact that, even when Congress identified circumstances warranting identical sentencing ranges regardless of drug quantity – for example, the twenty-to-life range when drug use causes death or serious bodily injury, or the mandatory life sentence when a prior felon traffics in drugs causing such injury – it repeated those penalties in each subsection rather than create a generally applicable provision.

Thus, if a prior drug felon, such as Gonzalez, traffics in an unquantified amount of crack, he faces a sentencing range of zero to thirty years. See id. § 841(b)(1)(C). But if such

a felon traffics in fifty grams or more of crack, that drug quantity increases his sentencing range to twenty years to life. See id. § 841(b)(1)(A). Nothing in the structure of the statute suggests that these corresponding minimums and maximums, or any of the others prescribed in the statute, can be delinked to permit mixing and matching across subsections to create hybrid sentencing ranges not specified by Congress. As the Ninth Circuit has observed, the government cannot “manipulat[e]” the sentencing scheme established by Congress in § 841 to “create[] a link where there is not one” between the minimum of one subsection and the maximum of another. United States v. Velasco-Heredia, 319 F.3d 1080, 1086 (9th Cir. 2003) (rejecting argument that § 841 mandatory minimum could apply without proof beyond a reasonable doubt of the statutory drug quantity). Much less does the statutory structure of § 841 suggest Congress’s intent to cast drug quantity in a dual role: performing as a sentencing factor for purposes of determining the applicable minimum sentence but as an element for purposes of determining the applicable maximum. Judge Becker made a similar point in his thoughtful concurring opinion in United States v. Vazquez, 271 F.3d 93, 113 (3d Cir. 2001) (en banc) (Becker, J., concurring): “It strains credulity . . . to assert that Congress intended for [drug] type and quantity to be treated as sentencing factors in some cases and as elements in others. I know of no statute written in such a manner, nor am I aware of any statutes construed this way.”⁷ Cf. United States v. Promise, 255 F.3d 150, 185 (4th Cir.

⁷ In Vazquez, the Third Circuit held that it was plain error to sentence a defendant to more than the § 841(b)(1)(C) twenty-year maximum when the jury had made no finding as to drug quantity. See United States v. Vazquez, 271 F.3d at 99-100. Nevertheless, it concluded that the error did not affect defendant’s substantial rights because the jury would undoubtedly have found the statutory drug quantity specified in § 841(b)(1)(B) proved

2001) (en banc) (Luttig, J., concurring in judgment) (“Either facts that affect the sentence a defendant receives are elements or they are not; they are not elements for some purposes and not for others.”).⁸

With this understanding of the structure of the statute at issue, we consider this court’s decision, after Apprendi v. New Jersey, 530 U.S. 466, to reconstrue the drug quantity provisions of § 841 as elements of an aggravated drug offense.

b. Our Post-Apprendi Identification of Drug Quantity as an Element of an Aggravated § 841 Offense

_____ At the time Gonzalez pleaded guilty to the charged conspiracy, this circuit, like every other to consider the question, had concluded that drug quantity was “a sentencing factor to be determined [by a preponderance of the evidence] by the district judge, not an element of _____
beyond a reasonable doubt. See id. at 100.

⁸ In Promise, the Fourth Circuit confronted an Apprendi challenge similar to that in Vazquez and found plain error that did affect defendant’s substantial rights but nonetheless declined to exercise its discretion to notice the error. See United States v. Promise, 233 F.3d at 152. In the view of a plurality of the court, “[t]here simply [could] be no doubt that had the indictment included the specific threshold quantity of 50 grams of cocaine base, the jury would have found Promise guilty beyond a reasonable doubt.” Id. at 164 (plurality opinion of Wilkins, J.). Concurring in the judgment, Judge Luttig concluded that the challenged sentence raised no Apprendi concerns if § 841 was properly construed to establish a maximum sentence of life imprisonment, id. at 169-70 (Luttig, J., concurring in judgment), in which case the statutory drug quantities operated as sentencing factors, id. at 173-74 (Luttig, J., concurring in judgment). We need not consider whether this “narrow” application of Apprendi, id. at 169 (Luttig, J., concurring in judgment), survives the Supreme Court’s decisions in Blakely v. Washington, 124 S. Ct. 2531 (2004), and United States v. Booker, 125 S. Ct. 738 (2005). As discussed in the next section of this opinion, our court has construed statutory drug quantity as an element of aggravated § 841 offenses. Nevertheless, we agree with Judge Luttig’s quoted observation that courts cannot sensibly construe a statute so that the same fact affecting sentences operates sometimes as an element and sometimes as a sentencing factor.

the offense to be proved by the prosecutor beyond a reasonable doubt and found by the jury.” United States v. Thomas, 274 F.3d at 662 (citing United States v. Thomas, 204 F.3d 381, 384 (2d Cir. 2000) (collecting cases from other circuits)); see also United States v. Boonphakdee, 40 F.3d 538, 542-43 (2d Cir. 1994); United States v. Monk, 15 F.3d 25, 27 (2d Cir. 1994); United States v. Campuzano, 905 F.2d 677, 679 (2d Cir. 1990). In so holding, we drew no distinction between drug quantities that defined statutory sentencing ranges and quantities that influenced the determination of Sentencing Guidelines ranges. See United States v. Moreno, 181 F.3d 206, 213-16 & n.4 (2d Cir. 1999).

The Supreme Court’s 2000 decision in Appendi v. New Jersey required us to reconsider our treatment of statutory drug quantities as sentencing factors. In Appendi, the Court ruled that,

[o]ther than the fact of a prior conviction, any fact that increases the penalty for a crime beyond the prescribed statutory maximum must be submitted to a jury, and proved beyond a reasonable doubt. With that exception, . . . “[i]t is unconstitutional for a legislature to remove from the jury the assessment of facts that increase the prescribed range of penalties to which a criminal defendant is exposed. It is equally clear that such facts must be established by proof beyond a reasonable doubt.”

530 U.S. at 490 (quoting Jones v. United States, 526 U.S. 227, 252-53 (1999) (Stevens, J., concurring)). The Court instructed that the relevant inquiry in determining whether a fact specified in the penalty provision of a criminal statute is an element of the crime that needs to be proved to a jury or a sentencing factor to be determined by the court is “one not of form, but of effect.” Id. at 494. If the fact “supports a specific sentence within the range authorized” by the jury verdict, then it can constitutionally be treated as a sentencing factor.

Id. at 494 n.19 (emphasis in original). On the other hand, if the fact would “expose the defendant to a greater punishment than that authorized by the jury’s guilty verdict,” then it must be deemed an element and submitted to the jury. Id. at 494.

Reconstruing the penalty provisions of § 841(b) in light of Apprendi, this court, sitting en banc in United States v. Thomas, 274 F.3d 655, unanimously concluded that the drug quantities specified in § 841(b)(1)(A) and -(b)(1)(B) were not mere sentencing factors but elements of aggravated offenses defined by reference to those statutes. In so holding, Thomas observed that “the principle of Apprendi requires quantity to be charged in the indictment and found by a jury only in cases where the quantity results in a punishment above a statutory maximum.” Id. at 660 n.3 (emphasis added). Nevertheless, Thomas did not reconstrue § 841 only with respect to prosecutions resulting in such sentences. Instead, it ruled that, “because the quantity of drugs involved in a crime may raise a defendant’s sentence above the statutory maximum established in 21 U.S.C. § 841(b)(1)(C), quantity is an element of the offense charged under 21 U.S.C. § 841.” Id. at 663 (emphasis added).⁹

This ruling illustrates the principle of constitutional avoidance: “‘When the validity of an act of the Congress is drawn in question, and . . . a serious doubt of constitutionality is raised, it is a cardinal principle that [courts] will first ascertain whether a construction of

⁹ In Thomas itself, because the defendant had not objected in the district court to being sentenced on the basis of a drug quantity that was not pleaded or proved to the jury, the court reviewed his Apprendi challenge for plain error, a standard of review that affords relief only if the defendant can demonstrate that the error affected substantial rights, as well as the fairness and public reputation of the judicial proceedings. See United States v. Thomas, 274 F.3d at 666-67. Because Gonzalez’s Apprendi/Thomas challenge was preserved in the district court, we need not here consider whether he satisfies these plain error requirements.

the statute is fairly possible by which the question may be avoided.” McConnell v. FEC, 540 U.S. 93, 180 (2003) (quoting Crowell v. Benson, 285 U.S. 22, 62 (1932)) (omission in original). By generally construing drug quantity as an element to be pleaded and proved in all prosecutions involving aggravated § 841 offenses, Thomas effectively ensured that no future § 841 conviction would be vulnerable to Apprendi attack.¹⁰

In United States v. Outen, this court reiterated Thomas’s general statutory construction of § 841, observing that drug quantity is properly construed as an element of offenses defined, in part, by § 841(b)(1)(A) or -(b)(1)(B) because, under those statutes, quantity “‘may be used to impose a sentence above the statutory maximum for an indeterminate quantity of drugs.’” 286 F.3d at 635 (quoting United States v. Thomas, 274 F.3d at 660). In such cases, “§ 841(a) no longer presents the entire offense; one element of the crime (drug quantity) is to be found in § 841(b).” Id. at 636 (emphasis in original).

This court has recognized the addition of a drug quantity element to a § 841(a) offense to result in a different criminal charge from the same offense pleaded without regard to

¹⁰ It is difficult to see how Thomas could have held otherwise without creating an intractable problem for parties and the courts. If the question of whether drug quantity is an element of an offense depended on the actual sentence imposed (specifically, on whether the sentence was above the otherwise applicable statutory maximum), it would be impossible to know until the final, sentencing phase of the litigation that drug quantity was an element of the crime of conviction that should have been pleaded in the indictment and proved beyond a reasonable doubt to the jury or admitted by the defendant.

In any event, as we discuss at greater length in Part II.B.2 of this opinion, *infra*, language in Apprendi and its Supreme Court progeny indicates that the Apprendi rule may sweep more broadly than suggested in Thomas, applying to facts that increase the applicable range of sentences and not simply to facts that actually result in sentences above otherwise applicable statutory maximums.

quantity. See United States v. Gaskin, 364 F.3d 438, 452-56 (2d Cir. 2004) (holding with reference to aggravated drug charges that “pleadings that add elements to the government’s burden of proof beyond those required for the lesser included charges . . . do more than gild the original charges. They replace them with different charges”) (citations omitted), cert. denied, 125 S. Ct. 1878 (2005); see also United States v. Thomas, 274 F.3d at 669-71 (noting that “lack of congruence” between unquantified drug offense proved and quantified drug offense on which defendant was sentenced might properly be analogized to a constructive amendment error). Accordingly, it has ruled that a defendant cannot be convicted on an aggravated offense unless the statutory drug quantity is proved to a jury or admitted by the defendant, see United States v. McLean, 287 F.3d at 134 (holding that where defendant “admits only the non-quantity elements” of a drug offense “and disputes quantity,” a district court may accept a guilty plea “only on the lesser-included offense of the § 841 crime involving an unspecified drug quantity” (emphasis in original)); United States v. Yu, 285 F.3d at 197 (holding it “error” to permit a defendant “to plead guilty to quantity-specific charges while refusing to allocute to quantity”); see also United States v. Doe, 297 F.3d 76, 90, 93 (2d Cir. 2002) (holding that plea allocution that did not settle issue of drug quantity required that defendant be sentenced pursuant to penalty provision applicable to indeterminate quantity offenses).

Thus, our statutory construction of § 841 – which applies generally to aggravated drug offenses, not simply those resulting in sentences raising Apprendi concerns – dictates that, in this case, Gonzalez could not be convicted of the charged § 841(b)(1)(A) conspiracy, or

sentenced pursuant to the increased penalty ranges provided in that subsection, unless a jury found or Gonzalez himself admitted the specified drug quantity element.

2. A Preponderance Finding as to Drug Quantity Does Not Require Imposition of a § 841(b)(1)(A) or -(b)(1)(B) Mandatory Minimum Sentence

The government does not dispute that Apprendi requires drug quantity to operate as an element in some cases involving aggravated § 841 offenses. Instead, it submits that Harris v. United States, 536 U.S. 545 (2002), United States v. Luciano, 311 F.3d 146 (2d Cir. 2002), and United States v. King, 345 F.3d 149 (2d Cir. 2003) (per curiam), establish that quantity is not an element when it triggers a § 841(b)(1)(A) mandatory minimum sentence (in this case, twenty years) that is not above the § 841(b)(a)(C) maximum that would apply to an identical unquantified drug offense (in this case, thirty years). Apparently, the same argument supported by the same cases prompted the district court to reconsider its initial inclination to grant Gonzalez’s motion to withdraw his guilty plea. This was unfortunate because these cases do not alter the fact that Thomas construed drug quantity as an element needing to be pleaded and proved beyond a reasonable doubt in every prosecution seeking conviction on an aggravated § 841 offense. See United States v. Thomas, 274 F.3d at 663 (“[B]ecause the quantity of drugs involved in a crime may raise a defendant’s sentence above the statutory maximum established in 21 U.S.C. § 841(b)(1)(C), quantity is an element of the offense charged under 21 U.S.C. § 841.” (emphasis added)). Thus, because Gonzalez did

not admit the drug quantity element of an aggravated § 841 offense, his guilty plea was insufficient to support conviction on a § 841(b)(1)(A) charge. Further, to the extent the cases cited by the government support imposition of mandatory minimum sentences on a preponderance of the evidence, we conclude, for reasons stated herein, that they do not control resolution of this appeal.

a. *Harris v. United States Does Not Pertain to Facts that Simultaneously Increase Mandatory Minimums and Maximums Thereby Creating an Increased Sentencing Range*

In Harris v. United States, the Supreme Court considered an Appendi challenge to a seven-year mandatory minimum sentence imposed pursuant to 18 U.S.C. § 924(c)(1)(A) for brandishing a firearm during and in relation to a drug-trafficking crime. See 536 U.S. at 551-52. That statute provides three increasing mandatory minimum sentences depending on a defendant's specific firearm use, while an implicit statutory maximum of life imprisonment remains constant throughout.¹¹ The Supreme Court ruled that firearm use was a sentencing

¹¹ The statute prescribes a five-year mandatory minimum sentence for any use or possession of a firearm in furtherance of a violent or drug crime, see 18 U.S.C. § 924(c)(1)(A)(i); a seven-year minimum if the firearm is “brandished,” id. § 924(c)(1)(A)(ii); and a ten-year mandatory minimum if the firearm is “discharged,” id. § 924(c)(1)(A)(iii). It identifies no explicit statutory maximum, implying a possible lifetime maximum, see United States v. Woodruff, 296 F.3d 1041, 1050 (11th Cir. 2002), although, as Justice Thomas notes in his Harris dissent, “virtually every . . . sentence imposed for a § 924(c)(1)(A) violation is tied directly to the applicable mandatory minimum,” Harris v. United States, 536 U.S. at 578, n.4 (Thomas, J., dissenting, joined by Stevens, Souter, and

factor rather than an element of the charged crime because facts “increasing the mandatory minimum (but not extending the sentence beyond the statutory maximum)” may properly be determined by a judge “without violating the Constitution.” Id. at 557 (plurality opinion of Kennedy, J.); see id. at 570 (Breyer, J., concurring in part and concurring in the judgment) (“I join [the Court’s] opinion to the extent that it holds that Apprendi does not apply to mandatory minimums.”). In so ruling, Harris reaffirmed McMillan v. Pennsylvania, 477 U.S. 79 (1986), a pre-Apprendi case holding that a court rather than a jury could determine whether a defendant visibly possessed a firearm during certain felonies because the aggravating fact only triggered a mandatory minimum sentence without altering the possible maximum. See Harris v. United States, 536 U.S. at 568.

The logic of the distinction drawn in Harris between facts that raise only mandatory minimums and those that raise statutory maximums is not easily grasped. See Harris v. United States, 536 U.S. at 569-70 (Breyer, J., concurring in part and concurring in judgment) (noting inability to “distinguish Apprendi v. New Jersey, 530 U.S. 466 (2000), from this case in terms of logic,” and expressing disagreement “with the plurality’s opinion insofar as it finds such a distinction”); see also id. at 579 (Thomas, J., dissenting, joined by Stevens, Souter, and Ginsburg, JJ.) (“Whether one raises the floor or raises the ceiling it is impossible

Ginsburg, JJ.).

to dispute that the defendant is exposed to greater punishment than is otherwise prescribed.”).

No matter. This case does not fit within the Harris paradigm.

As our earlier discussion of § 841 demonstrates, see supra at Part II.B.1.a, that statute, unlike 18 U.S.C. § 924(c)(1)(A), does not use a fact (drug quantity) simply to identify increasing minimum sentences within a penalty scheme with a fixed maximum. Instead, when drug quantity raises a mandatory minimum sentence under § 841, it simultaneously raises a corresponding maximum, thereby increasing a defendant’s authorized sentencing range above what it would have been if he had been convicted of an identical unquantified drug crime.¹² The Apprendi rule is, and after Harris remains, that “[i]t is unconstitutional for

¹² There is one exception: a change in crack quantity from five grams or more to fifty grams or more will increase a prior felon’s sentencing range from a prison term of ten years to life to a term of twenty years to life, compare 21 U.S.C. § 841(b)(1)(B) with id. § 841(b)(1)(A), an increase in only the minimum. Both specified quantities, however, increase a prior felon’s sentencing range above the zero-to-thirty year term provided in § 841(b)(1)(C) for unquantified crack offenses. We consider the exception insignificant to our analysis because it is unlikely that Congress intended drug quantity to operate as an element of an aggravated § 841 offense in all cases except when felons found guilty beyond a reasonable doubt of a -(b)(1)(B) quantity are sentenced to a § 841(b)(1)(A) mandatory minimum. See generally United States v. Vazquez, 271 F.3d at 114 (Becker, J., concurring) (“Although we can never be entirely certain of what Congress did intend in drafting § 841, we can be sure that Congress did not intend to design § 841 to fit precisely within the contours of a constitutional rule that did not yet exist.” (emphasis in original)). In any event, a quantity of crack that raises the minimum incarceratory sentence nonetheless always raises the overall § 841 sentencing range, because, in the case of a prior felon, a change from an unquantified amount of crack, to five or more grams, to fifty or more grams increases the possible fine from \$1,000,000 to \$2,000,000 to \$4,000,000 in the case of an individual,

a legislature to remove from the jury the assessment of facts that increase the prescribed range of penalties to which a criminal defendant is exposed.” United States v. Appendi, 530 U.S. at 490 (internal quotation marks omitted) (emphasis added). It would turn this rule on its head to conclude that a fact, such as drug quantity, which unquestionably increases the “range of penalties to which a criminal defendant is exposed,” is not an element of the crime that must be pleaded and proved beyond a reasonable doubt (or admitted by the defendant) because it increases the mandatory minimum sentence as well as the maximum. Harris simply does not speak to that circumstance. It holds that, where a fact “alter[s] only the minimum” sentence, without “authoriz[ing] the judge to impose [a] . . . higher penalt[y],” the constitution permits the fact to be treated as a sentencing factor. Harris v. United States, 536 U.S. at 554; cf. Jones v. United States, 526 U.S. at 242 (noting that McMillan v. Pennsylvania, 477 U.S. at 88, suggested that proof of visible gun possession might not have been deemed a sentencing factor if it had also “exposed a defendant to a sentence beyond the maximum that the statute otherwise set without reference to that fact”). That is not this case.

b. Drug Quantity Cannot Sensibly Be Construed to Operate as Either an Element or a Sentencing Factor Within the Same Aggravated Sentencing Range Depending on the Penalty Ultimately Imposed

The government submits that our decisions in United States v. Luciano, 311 F.3d 146,

unless he is a prior drug felon, in which case the increase is from \$2,000,000, to \$4,000,000, to \$8,000,000. Compare 21 U.S.C. § 841(b)(1)(C) with id. § 841(b)(1)(B), -(b)(1)(A).

and United States v. King, 345 F.3d 149, nevertheless, support the conclusion that drug quantity operates as a sentencing factor with respect to the mandatory minimum terms prescribed in § 841(b)(1)(A) or -(b)(1)(B) because those minimums never exceed otherwise applicable § 841(b)(1)(C) maximums. The argument is unconvincing on several levels.¹³

Preliminarily, we observe that the precedential value of Luciano and King is questionable. While Luciano did conclude that “the Apprendi rule does not apply to an increased mandatory minimum sentence unless the triggering circumstance results in a sentence in excess of an otherwise applicable statutory maximum,” 311 F.3d at 153, it did so in the context of reversing the grant of a writ of habeas corpus. Six months later, in Coleman v. United States, 329 F.3d at 90, this court ruled that Apprendi does not apply retroactively to habeas review. Thus, it appears that the reversal in Luciano should have been granted on this ground regardless of the merits of the petitioner’s Apprendi claim. To the extent United States v. King subsequently cited approvingly to Luciano in a direct appeal of a § 841 mandatory minimum sentence, that case’s discussion of the issue was dicta because the jury had, in fact, found the statutory drug quantity proved beyond a reasonable doubt.

¹³ As we noted at oral argument, if we were to adopt the position the government urges on this appeal its ability to secure guilty pleas from drug couriers and low-level participants in large-scale conspiracies might well be affected because a preponderance finding as to a statutory drug quantity would then require a district court to impose § 841(b)(1)(A) or -(b)(1)(B) mandatory minimum sentences in all § 841 cases, even when defendants pleaded guilty to lesser unquantified drug offenses.

See United States v. King, 345 F.3d at 151-52 (noting jury’s finding that offense involved five grams or more of crack, see 21 U.S.C. § 841(b)(1)(B)). The issue on appeal in King was whether Apprendi required the jury further to find that the defendant knew that the offense involved the statutory drug quantity, an argument at odds with our past precedent holding that the only knowledge required to prove an aggravated drug charge was that established in § 841(a). See id. at 151-53 (and cases cited therein).

In any event, developments in Apprendi jurisprudence suggest that the rule in that case may well reach more broadly than courts had originally understood.¹⁴ Luciano’s conclusion that the Apprendi rule does not apply to § 841 minimums was largely based on a determination that certain broad phrases – such as “[i]f a defendant faces” or if facts “expose the defendant to” greater punishment – used in Apprendi to identify circumstances requiring proof beyond a reasonable doubt to a jury were ambiguous. United States v. Luciano, 311 F.3d at 150-51 (quoting Apprendi v. New Jersey, 530 U.S. at 484, 494 (emphasis added)).

¹⁴ Certainly, the common assumption that Apprendi had no application to the federal Sentencing Guidelines, see United States v. Luciano, 311 F.3d at 153 (and cases cited therein); cf. United States v. King, 345 F.3d at 152 (relying on Thomas’s holding that quantity was not an element for purposes of Guidelines calculation, see United States v. Thomas, 274 F.3d at 663-64, to support conclusion that judicial factfinding supported statutory mandatory minimums), does not survive United States v. Booker, 125 S. Ct. 738. The possibility for still further change is suggested in Justice Thomas’s concurring opinion in Shepard v. United States, 125 S. Ct. 1254, 1264 (2005) (Thomas, J., concurring in part and concurring in judgment) (observing that a majority of the Supreme Court now views Almendarez-Torres v. United States, 523 U.S. 224 (1998) (upholding judicial factfinding as to defendant’s prior convictions), as inconsistent with Apprendi).

Luciano construed them to apply only when a defendant “is sentenced to a more onerous sentence than would have been allowed by law,” not when he is simply “exposed to the possibility of such a sentence,” but does not receive it. Id. at 151 (emphasis in original).

Since Apprendi, however, various pluralities of the Supreme Court, consisting of eight of the nine Justices, have persisted in using broad language, focusing on the increase in a sentencing range and not just an increase in the actual sentence, to identify facts that are properly treated as elements of aggravated crimes. In Harris, a plurality of the Court stated that “[t]hose facts that determine the maximum sentence the law allows,’ . . . are necessarily elements of the crime,” Harris v. United States, 536 U.S. at 565 (plurality opinion of Kennedy, J., joined by Rehnquist, C.J., O’Connor and Scalia, JJ.) (quoting Apprendi v. New Jersey, 530 U.S. at 499 (Scalia, J., concurring)) (emphasis added); and “those facts setting the outer limits of a sentence, and of the judicial power to impose it, are the elements of the crime for the purposes of the constitutional analysis,” id. at 567 (emphasis added).¹⁵ Similarly, in Sattazahn v. Pennsylvania, a plurality of the Court stated that Apprendi

¹⁵ As Luciano pointed out, the Harris plurality noted that “‘Apprendi said that any fact extending the defendant’s sentence beyond the maximum authorized by the jury’s verdict would have been considered . . . the domain of the jury by those who framed the Bill of Rights.’” United States v. Luciano, 311 F.3d at 152 (quoting Harris v. United States, 536 U.S. at 557 (plurality opinion of Kennedy, J.)) (emphasis and omission in Luciano). Nevertheless, that plurality went on to use the broader language quoted in the text of this opinion indicating that the Apprendi principle also reaches facts that “set[] the outer limits of a sentence.” Harris v. United States, 536 U.S. at 567 (plurality opinion of Kennedy, J.).

recognized as an element “any fact (other than a prior conviction) [that] increases the maximum punishment that may be imposed on a defendant, . . . no matter how the State labels it.” 537 U.S. 101, 111 (2003) (plurality opinion of Scalia, J., joined by Rehnquist, C.J., and Thomas, J.) (emphasis added). Most recently, in Shepard v. United States, a plurality of the Court alluded to the Appendi rule in just the terms that Luciano thought unlikely: “any fact other than a prior conviction sufficient to raise the limit of the possible federal sentence must be found by a jury.” 125 S. Ct. 1254, 1262 (2005) (plurality opinion of Souter, J., joined by Stevens, Scalia, and Ginsburg, JJ.) (emphasis added).

Thus, we cannot conclude, as the government urges, that Appendi and its progeny apply only to prosecutions that actually result in sentences exceeding otherwise applicable maximums. Nor can we conclude, after Appendi and our own decision in Thomas, that district courts are nevertheless required to treat drug quantity only as a sentencing factor for purposes of imposing § 841(b)(1)(A) and -(b)(1)(B) mandatory minimums. The Appendi rule applies to the resolution of any fact that would substitute an increased sentencing range for the one otherwise applicable to the case. Because mandatory minimums operate in tandem with increased maximums in § 841(b)(1)(A) and -(b)(1)(B) to create sentencing ranges that “raise the limit of the possible federal sentence,” Shepard v. United States, 125 S. Ct. at 1262, drug quantity must be deemed an element for all purposes relevant to the application of these increased ranges.

The Supreme Court’s decision in United States v. Booker, 125 S. Ct. 738 (2005), further supports this conclusion. Although Booker reiterates Apprendi’s rule by reference to the particular sentence imposed – “Any fact (other than a prior conviction) which is necessary to support a sentence exceeding the maximum authorized by the facts established by a plea of guilty or a jury verdict must be admitted by the defendant or proved to a jury beyond a reasonable doubt,” id. at 756 (emphasis added) – its ruling is not narrowly tailored to such cases but, instead, reaches every application of the federal Sentencing Guidelines. The Court explained that the Guidelines were suspect under Apprendi because they were “mandatory and impose[d] binding requirements on all sentencing judges” to sentence within prescribed ranges. Id. at 749-50. To remedy the problem, the Supreme Court did not deconstruct the Guidelines’ sentencing ranges, converting only the maximums to advisory provisions, while permitting the minimums to operate as mandatory sentencing factors. Instead, recognizing that Congress had structured the Guidelines as a unified system, the Court construed the whole as advisory to ensure against Apprendi error in particular cases.¹⁶

¹⁶ It appears that the Court recognized that not every Guidelines application implicated Apprendi concerns. See United States v. Booker, 125 S. Ct. at 753 (noting “many situations” in which a district court might find a Guidelines enhancement warranted, “yet still sentence the defendant within the range authorized by the jury”). The Court’s remedy, however, does not contemplate continued mandatory operation of the Guidelines in such circumstances. Rather, by severing and excising 18 U.S.C. § 3553(b)(1) and § 3742(e), the Court makes every application of the Guidelines advisory. Thus, while our court will not find a plain Booker error unless the sentence imposed on a defendant under a mandatory Guidelines scheme differs materially from the sentence that would have been imposed under advisory

So in Thomas, this court did not avoid Apprendi error in the application of increased sentences pursuant to § 841(b)(1)(A) and -(b)(1)(B) by construing quantity as an element only when a court sentences a defendant to a sentence higher than the maximum prescribed by the jury verdict or plea allocution; it construed drug quantity as an element in all prosecutions on aggravated charges. See United States v. Thomas, 274 F.3d at 663; see also United States v. Vazquez, 271 F.3d at 108 (Becker, J., concurring) (concluding that “drug type and quantity are always elements of an offense under § 841, and therefore must always be submitted to the jury for proof beyond a reasonable doubt” (emphasis in original)).

To the extent that United States v. King, 345 F.3d at 152, appears to have read Thomas more narrowly, we reiterate that this discussion was dicta and, in any event, relied on Thomas’s conclusion that it was unnecessary to treat drug quantity as an element in determining the (then-mandatory) Guidelines range applicable to a case, see United States v. Thomas, 274 F.3d at 663-64. We need not here consider the aptness of an analogy between the Guidelines and § 841 mandatory minimums because Booker no longer permits us to view a statutory mandate and an advisory guideline in the same light.

We recognize that some courts have been willing to cast drug quantity in a flexible

Guidelines, see United States v. Crosby, 397 F.3d 103, 118 (2d Cir. 2005), when presented with a preserved Apprendi challenge to the Guidelines, we have ordered resentencing without such a comparative analysis of the sentence imposed, see United States v. Fagans, 406 F.3d 138, 140-41 (2d Cir. 2005).

role with respect to aggravated § 841 prosecutions, with its identity as a sentencing factor or an element revealed only at sentencing. See United States v. Goodine, 326 F.3d 26, 31-32 (1st Cir. 2003); United States v. Copeland, 321 F.3d 582, 603 (6th Cir. 2003); United States v. Solis, 299 F.3d 420, 454 (5th Cir. 2002).¹⁷ But on this point, we are more inclined toward the views of the Fourth, Ninth, and District of Columbia Circuits, see United States v. Velasco-Heredia, 319 F.3d at 1085-87 (remanding for resentencing where court relied on judicial findings to impose mandatory minimum five-year sentence under § 841(b)(1)(B) that was not in excess of maximum for unquantified offense); United States v. Graham, 317 F.3d 262, 275 (D.C. Cir. 2003) (remanding for resentencing where jury verdict did not support imposition of five-year mandatory minimum term of supervised release pursuant to § 841(b)(1)(A), although same term could have been imposed under § 841(b)(1)(C) in court's discretion); see also United States v. Martinez, 277 F.3d 517, 530, 532-34 (4th Cir. 2002) (concluding that court erred in telling a defendant who was pleading guilty that he faced a mandatory minimum sentence of ten years and a maximum of life when he, in fact, faced no mandatory minimum sentence and a maximum of twenty years, although error did not affect substantial rights), as well as those expressed by Judge Becker in United States v. Vazquez, 271 F.3d at 108 (Becker, J., concurring) (reviewing structure and history of § 841 and

¹⁷ It appears that the Eighth and Tenth Circuits have also reached this conclusion in unpublished orders. See United States v. Sherman, 63 Fed. Appx. 279, 280 (8th Cir. 2003) (per curiam) (dicta); United States v. Ramirez, 43 Fed. Appx. 358, 362-63 (10th Cir. 2002).

specifically rejecting argument that drug type and quantity should be treated as “element-like factors only when they increase the prescribed statutory maximum penalty”).

Indeed, justice would hardly be well served by a rule that delayed the identification of one or more elements of a crime until sentencing. Prosecutors, who must draft indictments and develop evidence to meet their burden of proof; defendants and their counsel, who must decide whether to challenge the sufficiency of the government’s case or pursue plea negotiations; and trial judges, who must rule on the relevancy and sufficiency of evidence, prepare jury instructions, and ensure the factual bases for guilty pleas, all need to know long before sentencing which facts must be proved to a jury and which ones can be reserved for resolution by the sentencing judge. Thomas reconstrued § 841 to provide a clear answer to this question in light of concerns raised by Apprendi, and we repeat the court’s response today: drug quantity is an element that must always be pleaded and proved to a jury or admitted by a defendant to support conviction or sentence on an aggravated offense under § 841(b)(1)(A) or -(b)(1)(B). If a defendant is convicted only on a lesser unquantified drug charge, he must be sentenced pursuant to § 841(b)(1)(C), which generally provides no mandatory minimum sentence.

In sum, while the district court in its discretion could have sentenced Gonzalez to a twenty-year term of incarceration pursuant to § 841(b)(1)(C), unless a jury found or Gonzalez admitted the charged statutory drug quantity, the court was not required to sentence him to

that term pursuant to § 841(b)(1)(A), nor was its departure discretion curbed by that mandatory minimum.

C. Gonzalez’s Plea to a § 841(b)(1)(A) Offense Cannot Be Deemed Knowing, Voluntary, or Sufficient

_____ 1. The Rule 11(b)(1)(C) Defect

Although district courts enjoy considerable discretion in ruling on motions to withdraw guilty pleas, this court has, nevertheless, required withdrawal to be granted when a defendant was not provided with the “mix of information” required by Rule 11 unless, of course, the error can be deemed harmless because it would not have affected the defendant’s decision to plead guilty. United States v. Harrington, 354 F.3d at 183-84; see also Fed. R. Crim. P. 11(h) (“A variance from the requirements of this rule is harmless error if it does not affect substantial rights.”).

In Harrington, a defendant who, like Gonzalez, pleaded guilty before Appendi/Thomas to federal drug charges, was told by the court that he faced a mandatory minimum sentence of twenty years and a maximum sentence of life imprisonment, although the drug quantity specified in § 841(b)(1)(A) for this sentencing range was neither pleaded in the indictment nor explicitly admitted by the defendant in the course of his plea allocution. See 354 F.3d at 180, 181, 184. Finding the defendant to have been misinformed as to the applicable minimum and maximum sentences in his case, see id. at 184 (“[U]nder Appendi, the actual sentence range on the two counts to which [defendant] pled guilty was zero-to-

thirty years' imprisonment.”), this court ruled that the defendant's plea violated Fed. R. Crim. P. 11(b)(1)(H) (requiring court to “inform the defendant of, and determine that the defendant understands . . . any maximum possible penalty, including imprisonment”), and 11(b)(1)(I) (stating same requirement with respect to “any mandatory minimum penalty”), and that the error could not be deemed harmless, see id. at 184-86. The court explained that the erroneous information about the applicable mandatory minimum was particularly troublesome because “a mandatory minimum sentence represents such a strong inducement to plea” that misinformation as to the applicability of such a minimum “is presumptively significant in the defendant's decision-making.” Id. at 185-86. Because of that misinformation and because the record revealed “sufficient confusion” as to whether the defendant admitted or sought to challenge the quantity of drugs involved in the charged conspiracy, the court found that the Rule 11 error was not harmless and remanded with instructions to allow the defendant to withdraw his plea. Id. at 186 (internal quotation marks omitted).

Following Harrington,¹⁸ we conclude that Gonzalez was also misinformed about the

¹⁸ The government submits that Harrington cannot be reconciled with this court's prior rulings in Luciano and King. For reasons already discussed, we conclude that the discussion of § 841 mandatory minimums in the latter two cases is entitled to little weight. Because this decision might, nevertheless, be seen as a departure from those earlier ones, we have circulated this opinion before filing to all active members of this court. See, e.g., United States v. Crosby, 397 F.3d at 105 n.1 (2d Cir. 2005); United States v. Mincey, 380 F.3d 102, 103 n.1 (2d Cir. 2004) (per curiam).

operation of § 841(b)(1)(A)’s mandatory minimum and maximum sentences, resulting in a violation of Rule 11: as in Harrington, we conclude that the circumstances surrounding the plea show that the error was not harmless. To be sure, this case differs from Harrington in that the quantity of crack cocaine necessary to apply § 841(b)(1)(A)’s twenty-years-to-life sentencing range (rather than § 841(b)(1)(C)’s zero-to-thirty year range) was pleaded in the indictment. Thus, because the district court correctly informed Gonzalez that the charged crime exposed him to a sentence within this range, there was no Rule 11(b)(1)(H) or - (b)(1)(I) violation in this case. Instead, there was an error with respect to Rule 11(b)(1)(C), which requires a court to ensure that a defendant understands his “right to a jury trial” on the charge to which he is pleading guilty.

Not surprisingly given that Gonzalez’s guilty plea pre-dated Apprendi and Thomas, the district court did not inform him of his right to have the pleaded statutory drug quantity proved to a jury. Instead, the court advised Gonzalez that its own preponderance finding as to that drug quantity was all that was necessary to impose sentence pursuant to § 841(b)(1)(A)’s twenty-year mandatory minimum and its possible lifetime maximum. For reasons already discussed, we conclude that, after Thomas, this was an incorrect statement of the law. Thus, in this case, no less than in Harrington, the Rule 11 violation deprived defendant of information that was significant to his decision-making as to whether to plead guilty. Nor is there any indication that the presumptively significant information regarding

the operation of a mandatory minimum sentence was in fact insignificant in the particular circumstances of this case. The record plainly demonstrates that the only fact Gonzalez wished to contest with respect to the charged conspiracy was drug quantity precisely because of its potential impact on any term of incarceration, both in terms of the statutory minimum and the Sentencing Guidelines. If the § 841(b)(1)(A) drug quantity was an issue to be resolved by the judge, as Gonzalez was mistakenly informed, there was no reason for him to proceed to trial; better to plead guilty and concentrate his efforts on disputing quantity at a Fatico hearing. But if Gonzalez had understood that the issue was one that the government would have to prove to a jury at trial, it is doubtful that he would have waived that right. See generally Plea Tr. at 12-13 (defense counsel notes that, because defendant acknowledged conspiring to distribute a sample of crack and because the disputed issue of a larger quantity would be resolved by the court, there was no purpose in proceeding to trial; defendant's "only option" was to plead guilty and concentrate his efforts on the Fatico hearing). Indeed, once the matter had been put to a hearing, the district court found that the evidence supported a finding of drug quantity sufficient to trigger § 841(b)(1)(A) by a preponderance but not necessarily by a reasonable doubt, demonstrating that Gonzalez's reluctance to plead to any such amount was strategically sound. Thus, after Apprendi/Thomas, the district court could not conclude that Gonzalez knowingly and voluntarily pleaded guilty to an aggravated drug offense under § 841(b)(1)(A) based on misinformation that a judicial finding as to drug

quantity, rather than jury verdict or plea, would mandate imposition of that statute's minimum twenty-year sentence. Accordingly, the motion to withdraw the guilty plea should have been granted. See United States v. Harrington, 354 F.3d at 186.

2. The Rule 11(b)(3) Defect

As a precondition to entering judgment, a court must “determine that there is a factual basis for the plea.” Fed. R. Crim. P. 11(b)(3). The statutory entries in Gonzalez’s judgment reflect his conviction for a conspiracy, see 21 U.S.C. § 846, to traffic in crack cocaine, see id. § 812, in violation of § 841(a)(1) and -(b)(1)(A). As we observed in United States v. Outen, “in such a situation § 841(a) no longer presents the entire offense; one element of the crime (drug quantity) is to be found in § 841(b)[(1)(A)] instead.” 286 F.3d at 636 (emphasis in original).

For reasons already discussed, we conclude that a drug conspiracy pleaded by reference to § 841(b)(1)(A) is a different, aggravated offense from an unquantified drug conspiracy, see United States v. Thomas, 274 F.3d at 669-71; see also United States v. Gaskin, 364 F.3d at 452-56, and that Gonzalez’s failure to admit – indeed, his explicit challenge to – the statutorily prescribed quantity meant that his plea did not provide the court with an adequate factual basis to enter judgment against him on the charged crime, see United States v. McLean, 287 F.3d at 133-34; United States v. Yu, 285 F.3d at 197-98; see also United States v. DeJesus-Abad, 263 F.3d 5, 9 (2d Cir. 2001) (per curiam) (noting that

court must assure itself that conduct admitted by defendant is “an offense under the statutory provision under which he is pleading guilty” (internal quotation marks omitted)). Accordingly, this defect provided a further meritorious ground for granting Gonzalez’s motion to withdraw his guilty plea to a conspiracy defined by reference to § 841(b)(1)(A).

III. Conclusion

To summarize, we conclude:

- (1) The drug quantities specified in 21 U.S.C. § 841 are elements that must be pleaded and proved to a jury or admitted by a defendant to support any conviction on an aggravated drug offense, not simply those resulting in sentences that exceed the maximum otherwise applicable for an identical unquantified drug crime.
- (2) The sentencing ranges prescribed in § 841 for aggravated drug offenses may not be deconstructed so that quantity operates as an element for purposes of determining an applicable maximum but as a sentencing factor for purposes of determining an applicable minimum. Thus, where a drug quantity specified in § 841(b)(1)(A) or -(b)(1)(B) is neither proved to a jury nor admitted by a defendant, a district court is not required to impose the minimum sentence mandated by those sections even if it may impose that same sentence pursuant

to § 841(b)(1)(C).

- (3) Because the defendant in this case was misinformed as to his right to have the charged statutory drug quantity proved to a jury and because he did not admit quantity at his plea allocution, his guilty plea to an aggravated § 841(b)(1)(A) offense was not knowing, voluntary, or sufficient to support the judgment of conviction. The circumstances of this case do not show that Gonzalez would have pleaded guilty to the offense had he been properly advised; thus, the error was not harmless. Further, because the government was unwilling to accept defendant's plea to an unquantified drug offense in satisfaction of the charge, the defendant should have been allowed to withdraw his guilty plea.

The case is REMANDED with directions to vacate the judgment of conviction, to allow the defendant to withdraw his guilty plea, and to permit the government to proceed to trial on the aggravated charge.