

# Supplementary Material for SPoRC-VIST: A Benchmark for Evaluating Generative Natural Narrative in Vision-Language Models

Yunlin Zeng  
Georgia Institute of Technology  
yzeng@gatech.edu

## A. Inference Prompt

For all inference experiments, we provide the model with the following system prompt along with the five input images:

*“Generate a natural conversational podcast dialogue. Use the format Speaker 1:, Speaker 2:, Speaker 3:, etc. for multiple speakers. Do not reference the images or use phrases like ‘our first image’. Write casual, authentic spoken dialogue without introductions or sign-offs. The word count should be around 800 words.”*

This prompt was designed to elicit natural podcast-style dialogue while preventing the model from producing meta-commentary about the images themselves.

## B. Full Transcript Comparison

Here we present the full, unedited transcripts for both evaluation examples shown in Figure 1 of the main paper, as generated by the 32B Base, 32B Fine-tuned (Ours), and 235B Base models. Sentences that directly reference visual content from the input images are highlighted in red text, with the corresponding image index noted in parentheses.

### Example 1: Wedding Party (Story ID: 50242)

This corresponds to Figure 1, Row 1 in the main paper.

#### B.1. Qwen3-VL 32B Base Model (689 words)

**Alex:** Hey, Jamie! You see these photos I sent over? I was going through some old wedding pics from last year, and I just got hit with a wave of nostalgia. I mean, look at this first one—five guys in suits, standing outside some grand old building, laughing. (Image 1) It’s like they’re plotting something. Or maybe just sharing a joke before the ceremony.

**Jamie:** Oh, I love that one! There’s something so authentic about it. You can tell they’re not posing—they’re just there, in the moment. And that guy with the camera

bag? (Image 1) He’s probably the photographer, right? But he’s laughing too, like he’s part of the crew, not just the observer. I think that’s what I love about wedding photos—they capture the real, unscripted stuff.

**Alex:** Exactly. And then you see the third photo—the bride and groom outside, surrounded by friends. She’s got that little net veil, and he’s in a grey vest with his sleeves rolled up. (Image 3) It’s so relaxed, so modern, but still elegant. And they’re all holding those blue plastic cups—I mean, come on, that’s such a fun detail. (Image 3) It’s not fancy champagne flutes; it’s real celebration.

**Jamie:** Oh, I love that! It’s like, “We’re not pretending to be fancy. We’re just happy.” And the greenery behind them—it’s lush, natural. It feels like they’re in a garden, maybe a courtyard. (Image 3) It’s intimate, not staged.

**Alex:** And then you get to the cupcakes—the fifth photo. Oh my god, look at those! They’re on these elegant tiered stands, but they’re cupcakes. (Image 5) Not a big cake. And they’re all different—chocolate, vanilla, pink frosting, swirls. (Image 5) It’s like a dessert buffet.

**Jamie:** I can! It’s like the perfect end to the day—sweet, colorful, and shared. And the fact that they chose cupcakes instead of a traditional cake? That’s such a modern touch.

**Alex:** Exactly. And I think that’s the theme that ties all these photos together—authenticity. Whether it’s the guys laughing outside the building (Image 1), the bride in her simple dress, the chaotic reception (Image 4), or the cupcakes (Image 5)—it’s all about being true to who you are.

**Jamie:** Well said. Here’s to messy, beautiful, real moments—and to cupcakes.

**Alex:** Cheers to that.

#### B.2. Qwen3-VL 32B Fine-tuned (Ours) (1,191 words)

**Speaker 1:** So I was thinking, what if we did a podcast about weddings? Not like, the whole “how to plan your dream day” kind of thing, but more like, the real stories behind them. The weird stuff, the funny moments, the things you’d never tell your grandma.

**Speaker 2:** Oh, I’m in. I’ve been to enough weddings

to know that 90% of them are just people pretending to be happy while someone else's life is being celebrated. But the other 10%? That's gold. Like that time I saw a guy in a full suit trying to sneak a sandwich into the ceremony. He was like, "I'm not hungry, I'm just... carrying it for my mom."

**Speaker 3:** Hah! That's classic. I remember one wedding where the best man got so drunk during the speeches he started quoting Shakespeare, but he kept mixing up Hamlet and Macbeth. He was like, "To be or not to be... wait, no, that's the one with the witches, right?" The whole room was silent. Then he looked at the bride and said, "You're my Juliet, and I'm your... uh... Romeo? Or was it Mercutio? I'm confused."

**Speaker 1:** I've got one that's even better. **My cousin got married at this super fancy old courthouse. (Image 2)** Everyone was dressed up, the whole thing was formal. But then, halfway through the ceremony, **this guy walks in with a giant camera and a backpack, starts snapping photos like he's covering a war zone. (Image 1)** And he's not even part of the wedding party.

**Speaker 2:** Wait, wait, wait—did he get kicked out?

**Speaker 1:** No! The bride's brother actually invited him. Turns out he's a freelance photographer and he'd been following the couple for months, documenting their "love story" without them knowing. The bride was like, "Wait, you've been stalking us?" And he was like, "Stalking? I was *documenting!*"

**Speaker 3:** That's insane. But also kind of sweet, in a creepy way.

**Speaker 2:** Speaking of weird, I once went to a wedding where the reception was in a barn, but **they had a full-on cupcake tower. Like, three tiers, 50 cupcakes, all different flavors. (Image 5)** And someone had clearly forgotten to tell the caterer that the bride was allergic to nuts. So when they cut into the first one, it was like, "Oh no, it's got almonds in it!" And the bride's face turned white. She didn't say anything, just handed it to the groom and said, "You eat it. I'm not risking anaphylaxis for a wedding cake."

**Speaker 1:** That's terrifying. But also, I love that she didn't panic. She just handed it off like, "You're the one who said 'I do'—you deal with it."

**Speaker 3:** I've got a story that's more about the aftermath. My friend got married in this tiny village hall, and after the ceremony, they had a little reception with sandwiches and wine. But the thing was, **the groom's dad had brought his laptop. Not for work. For the reception. He was sitting at a table, typing away, completely ignoring everyone. (Image 4)** We were like, "Dad, are you working?" And he was like, "No, I'm just... updating my LinkedIn. You never know when you'll need it."

**Speaker 2:** That's the most relatable thing I've ever heard.

**Speaker 1:** **And then there's the one where the bride**

**wore a veil, but it was like a tiny birdcage thing. And she was standing next to the groom, who was in a vest, (Image 3)** and they were both looking at the guests like, "Is this happening?" And then she took a sip of water and the veil got caught on the rim of the glass.

**Speaker 3:** I've seen that. I've seen people trip on their train, spill wine on the groom, forget their vows... it's chaos. But the best part is, no one really cares. Everyone's just happy to be there. That's what makes weddings so beautiful. It's not perfect. It's human.

**Speaker 2:** Exactly. I think that's why I love them. It's not about the dress or the venue or the cake. It's about the people.

**Speaker 1:** So we're doing it? The "Weird Weddings" podcast?

**Speaker 2:** I'm in. Let's call it "The Unplanned Vows."

**Speaker 3:** I like it. And we can have a segment called "Cupcake Catastrophes."

**Speaker 1:** Perfect. We'll start with the courthouse photographer story. That's a solid opener.

### B.3. Qwen3-VL 235B Base Model (979 words)

**Speaker 1:** Alright, so I'm still trying to process how we ended up at a wedding in a courtroom. Like, seriously? Was that even legal? **I mean, the wood paneling, the chandeliers, the judge's bench with the green curtain behind it (Image 2)**—it felt like we were about to get sentenced, not celebrate.

**Speaker 2:** It was totally legal! The registrar was legit. And honestly? It was kind of brilliant. No boring church, no awkward hymns—just a really elegant room with a weirdly regal vibe. Plus, the lighting? Chef's kiss. You could see every detail—the flowers, the hats, **even the guy in the front row trying to sneak a selfie with his phone. (Image 2)**

**Speaker 3:** That guy was me. And I got a solid shot of the bride walking in. She looked amazing. **That veil, the dress—it was like she stepped out of a vintage magazine. (Image 3)** And the groom? He was beaming. You could tell he was trying not to cry, but his eyes were definitely glistening.

**Speaker 4:** I think everyone was trying not to cry. **Even the guy with the mustache who was sitting next to me—he was wiping his eyes with his napkin. And then he started recording the whole thing on his phone like he was making a documentary. (Image 2)** I swear, I thought he was going to stand up and yell "Objection!" halfway through.

**Speaker 1:** Honestly, I think that would've been the highlight. "Objection! The groom is too handsome!" But seriously, the whole thing was so... intentional. Like, they didn't just pick a venue—they curated an experience. Even the reception was a vibe. **That checkered tablecloth? The Union Jack bunting? (Image 4)** It was like a British garden party meets a hipster wedding.

**Speaker 2:** And don't even get me started on the food. I mean, I was expecting a few sandwiches and some cheap wine, but no—they had this whole spread. Cheese boards, little sandwiches, and then the cupcakes. (Image 4, 5) Oh my god, the cupcakes. They were stacked on these tiered stands like they were in a fancy bakery window. (Image 5) Chocolate, vanilla, pink swirls (Image 5)—I think I had three.

**Speaker 3:** You had five. I counted. And you were trying to be sneaky about it, but I saw you sneak back for that chocolate one with the pink frosting. You're lucky the bride didn't catch you—she was busy trying to keep her veil from getting caught in her chair.

**Speaker 4:** That veil was a menace. I saw her trying to adjust it while she was talking to the guests. And then there was that moment when she was standing by the window, and the light hit her just right—it was like a movie scene. I actually stopped eating for a second to take it in.

**Speaker 1:** Same. I think we all did. It was one of those moments where you just pause and think, “This is why we do this.” You know? Weddings are weird—they're this mix of chaos and beauty, and somehow it all comes together.

**Speaker 2:** And let's not forget the groom's speech. I mean, he started off with a joke about how he used to think marriage was just a legal contract, and then he got all serious and said, “But now I know it's about choosing to be happy with the same person every day.” I swear, I choked on my wine.

**Speaker 3:** I think half the room choked on their wine. And then the best man had to step in and say, “Alright, enough of that—let's get back to the cake!”

**Speaker 4:** I was waiting for someone to accidentally stab themselves with it. But nope—everyone was surprisingly coordinated. Even the little girl in the pink dress who was sitting next to me—she was eating her cake like a pro. No mess, no spills, just pure focus.

**Speaker 1:** That kid was a legend. I think she's going to be a wedding planner one day.

**Speaker 2:** You did spill your drink on your suit. Twice. But hey, at least you looked good doing it. And honestly, that's what weddings are about—looking good while you're making a mess. It's like a rite of passage.

**Speaker 3:** I think the real rite of passage was surviving the group photo. (Image 1) I swear, we must've taken 20 pictures before everyone was smiling at the same time.

**Speaker 4:** But we pulled it off. And now we have these amazing photos to prove it. I mean, I've already sent them to my mom, and she's already planning her own wedding based on this one. She wants the courtroom, the cupcakes (Image 5), the checkered tablecloths—the whole shebang.

**Speaker 1:** Tell her to add the mustache guy with the phone. (Image 2) He's a must-have.

**Speaker 2:** Agreed. And hey, if we're lucky, maybe

we'll get invited to the next one. I hear they're thinking about doing it on a boat. With fireworks. And more cupcakes.

**Speaker 3:** I'm already packing my bag.

## Example 2: Motorcycle & Marathon (Story ID: 47306)

This corresponds to Figure 1, Row 2 in the main paper.

### B.4. Qwen3-VL 32B Base Model (689 words)

**Alex:** Hey, Jamie! You see those photos I sent over? The ones with the woman in the leather gear (Image 1) and then the marathon stuff? I was flipping through them and I just had to talk about them. There's something really cool about how they tell a story—not just of one person, but of different sides of life, you know?

**Jamie:** Oh, I saw them! I was actually kind of struck by the contrast. Like, first you've got her in this full-on black leather jacket and pants, standing in a room, holding a helmet (Image 1), smiling like she's ready to take on the world. And then—boom—she's on a motorcycle, helmet on, looking like she's about to ride off into the sunset. (Image 2) And the bike—it's a Yamaha, right? Not some cruiser. It's sporty, sleek. (Image 2) She's not just riding for fun; she's riding with purpose.

**Alex:** And then—cut—she's at a marathon. And she's not even in the same outfit! She's in this soft white turtle-neck, which is so different from the leather. (Image 3) It's almost like she's shedding that tough exterior for something more... approachable? Or maybe just more comfortable for running. But she's still smiling. Still confident. Just in a different way.

**Jamie:** Right? And the setting changes too. From a quiet room, to a parking lot with a car and a hedge (Image 2), to this huge, crowded grassy field with a giant “Íslandsbanka Reykjavík Maraton” banner. (Image 3) It's like she's moving from private to public, from solitary to community.

**Jamie:** It's like the marathon becomes this big, messy, beautiful human moment. And she's part of it—not just as a runner, but as someone who's *feeling* it. You can see it in her face in the first marathon photo—she's not just posing. She's *there*. She's present. And then, in the running shot, she's laughing. (Image 4, 5) That's the real win, isn't it? Not finishing first, but finishing with joy.

**Alex:** Absolutely. And I think that's the thread connecting all the photos—this sense of *living fully*. Whether she's riding a motorcycle (Image 2), standing in her leather gear (Image 1), or running a marathon (Image 3, 4, 5), she's engaged. She's not just going through the motions. She's choosing to be in the moment, to embrace the challenge, to wear the gear that fits the moment.

**Jamie:** And maybe, just maybe, to laugh while you're doing it.

**Alex:** Always. Always laugh.

### B.5. Qwen3-VL 32B Fine-tuned (Ours) (935 words)

**Speaker 1:** So I was looking through some old photos the other day, and I came across these pictures from a few years back. I was trying to remember what the heck I was doing, and then it all came back. I had just gotten my motorcycle license, and I was so excited. **I remember buying this black leather jacket and matching pants. (Image 1)** I thought I looked like a total badass.

**Speaker 2:** That’s awesome! I remember you showing me the jacket. You were so proud of it. **You even had the whole outfit—the boots, the gloves, the helmet. (Image 1)** You looked like you were ready to take on the open road.

**Speaker 1:** I was! **I had this little Yamaha (Image 2)**, and I’d ride around town just to feel the wind in my face. It was such a rush. I even did a little stunt once—just a tiny wheelie—and I almost wiped out. I laughed so hard afterward. I was more scared than I let on, but I felt so alive.

**Speaker 2:** You always did love that adrenaline, didn’t you? I remember you telling me you wanted to ride across Europe someday. You even had a map on your wall with all the routes you wanted to take.

**Speaker 1:** I did! I still have that map. I didn’t end up doing it, but I’ve been thinking about it lately. Maybe next year? I mean, I’m not getting any younger, and I don’t want to wait too long.

**Speaker 3:** You know, I think you should. Life’s too short to wait for the “perfect” time. You’ve always been the kind of person who just goes for it. I remember when you signed up for that marathon—out of nowhere—and you didn’t even train for it. You just showed up and ran it.

**Speaker 1:** Ha! Yeah, that was a disaster. I think I ran the whole thing at a jog, and I was so sore for a week. But you know what? I loved it. I remember the day of the race—it was raining, and **the streets were wet (Image 5)**, but there was this energy in the air. Everyone was so excited. **I was wearing this white jacket (Image 3)**, and I felt like I was part of something big.

**Speaker 2:** You were! I remember you texting me during the race, saying, “I can’t believe I’m doing this!” I was so proud of you. You didn’t even train, and you finished. That’s impressive.

**Speaker 3:** And you didn’t just finish—you **ran with this huge smile on your face the whole time. I saw you in the photos, and you looked like you were having the time of your life. (Image 3, 5)** I think that’s what I love most about you—you find joy in the little things, even when they’re challenging.

**Speaker 1:** Well, I guess I’m just a sucker for a good challenge. I mean, I didn’t even know I liked running until I signed up for that race. I thought I’d hate it. But something

about being out there, surrounded by people who were all pushing themselves—it just felt right.

**Speaker 2:** It’s funny how sometimes we do things on a whim, and they end up changing us. I think that’s what happened with you and the motorcycle, too. You weren’t really into it before, but then you got your license, and suddenly it was your thing.

**Speaker 1:** Exactly. I think we all have those moments—that one thing that just clicks. For me, it was the freedom of the road and the rhythm of running. Both of them made me feel like I was truly living.

**Speaker 3:** You know, I think that’s why you’re such an inspiration to people. You don’t wait for permission to do the things you want. You just go for it. And that’s something a lot of us could learn from.

**Speaker 1:** Well, I didn’t plan any of it. I just followed my curiosity. I remember thinking, “What if I try this?” and then I did. And most of the time, it turned out okay—sometimes even better than I expected.

**Speaker 2:** Like when you rode that motorcycle through the rain and got soaked, but you still grinned the whole way home.

**Speaker 1:** (laughs) I did! I was covered in mud, and my jacket was dripping, but I didn’t care. I felt like I’d conquered the world. I think that’s what it’s all about—not being afraid to get messy, to take risks, to just... be.

**Speaker 3:** I think that’s the real takeaway here. Life isn’t about being perfect or having everything figured out. It’s about showing up, even when you’re scared, and just doing it. And you’ve done that more than most.

**Speaker 1:** Thanks, I guess. I just try to live in the moment. And hey, if I can inspire even one person to try something new—whether it’s riding a motorcycle or running a marathon—then I’ve done something right.

**Speaker 2:** You’ve done more than that. You’ve reminded us all that we’re capable of more than we think. And that’s a pretty big deal.

**Speaker 3:** Absolutely. So here’s to you—to the adventures, the mishaps, the rain-soaked runs, and the muddy rides. May you keep going, keep exploring, and keep smiling through it all.

**Speaker 1:** (smiling) I’ll try my best. And who knows? Maybe next year, I’ll be riding across Europe with that map finally coming to life.

### B.6. Qwen3-VL 235B Base Model (910 words)

**Speaker 1:** You know what’s wild? I used to think I’d never be the kind of person who’d **wear head-to-toe leather (Image 1)** just to look cool. But then I got into motorcycling, and suddenly, it’s not about looking cool—it’s about not getting scraped off the pavement. **That jacket? Saved my ass more than once. (Image 1)**

**Speaker 2:** Wait, you’re the one who rides that silver

Yamaha? (Image 2) I remember seeing you parked next to that old blue truck. (Image 2) You looked like you were about to take off for Reykjavík or something. Did you actually go anywhere epic?

**Speaker 1:** Ha! No, just to the grocery store. But I swear, riding through the city in full gear, people stare like you're some kind of outlaw. Honestly, I just want to get my oat milk without getting rained on. The helmet's the real MVP—it keeps the wind out of my ears and the bugs out of my mouth.

**Speaker 3:** Okay, but let's talk about the marathon. I saw you at the start line, all chill in that white fleece (Image 3), sunglasses on, like you were just out for a Sunday stroll. And then—bam—you're sprinting past people like you've got a personal vendetta against the finish line.

**Speaker 1:** That was the Reykjavík Marathon, right? (Image 3) Yeah, I was feeling good that day. The crowd was insane—everyone cheering, waving, some guy even handed me a beer as I ran by. I didn't take it, obviously, but the gesture? Priceless. You should've seen the look on the guy's face when I waved and kept going.

**Speaker 2:** I was there too! I was one of the idiots running in the blue shirt with the red cap. (Image 4) You passed me right after the hill near the park. I was wheezing like a broken accordion. How do you even train for that? Do you just... run everywhere?

**Speaker 1:** Pretty much. I run to the store, to work, to the bus stop—I even ran to my own birthday party once. My friends thought I was nuts. But honestly, once you get into it, it's addictive. The rhythm, the sweat, the way your legs just... keep going. Plus, it's a great way to see the city. You notice stuff you'd never see in a car.

**Speaker 3:** I tried running once. Made it two blocks before I had to stop and pretend I was looking at a map. I think my lungs are still mad at me. But hey, at least I can appreciate the view from the sidelines. Like, watching you guys run past those green buildings? That was a vibe. The crowd was so hyped—people were dancing, clapping, some even had signs. (Image 5) One guy had a sign that said “You're doing great, even if you're last!” I cried a little.

**Speaker 1:** That's the spirit! Marathons aren't just about speed—they're about showing up, putting one foot in front of the other, and not letting the rain or the hills or the guy with the beer stop you. I mean, I've had days where I wanted to quit halfway, but then I'd see someone cheering in a ridiculous costume, and I'd be like, “Okay, I can do this for five more minutes.”

**Speaker 2:** I love that. It's like, life's a marathon, right? You don't have to be the fastest, just the most stubborn. Speaking of stubborn—how's the motorcycle thing going? Still riding in all weather?

**Speaker 1:** Absolutely. Rain, snow, sleet—I've done it all. There's something about the wind hitting your face, the

engine rumbling under you, that just clears your head. Plus, it's a great way to escape the city when you need a break. I'll just hop on, ride out to the coast, and watch the waves crash. It's my therapy.

**Speaker 3:** That sounds amazing. I'm more of a couch-and-coffee kind of person, but I get it. Sometimes you just need to move, you know? Whether it's on two wheels or two feet, it's about feeling alive. And hey, if you ever need a running buddy, I'll try not to slow you down too much.

**Speaker 1:** Deal. But fair warning—I'm not slowing down for anyone. You'll have to keep up. And if you're lucky, I might even let you borrow my helmet. Just don't scratch it.

**Speaker 2:** Oh, I'll take my chances. But seriously, you're inspiring. I think I'm going to start running again. Maybe not a marathon—baby steps, right? Like, a 5K. Or even just to the corner store.

**Speaker 1:** That's the spirit! Start small, build up. And hey, if you ever want to ride shotgun on the bike, just say the word. I'll even let you pick the playlist. As long as it's not country.

**Speaker 3:** Country? No way. I'm more of a punk rock kind of person. But I'll settle for some good indie vibes. And maybe a post-run coffee? I know this place downtown that has the best oat milk lattes.

**Speaker 1:** Sold. Let's make it a thing. Run, ride, repeat. And who knows? Maybe next year, we'll all be at the marathon together. You in the blue shirt (Image 4), me in the white fleece (Image 3), and you, Speaker 2, in whatever ridiculous outfit you can find.

**Speaker 2:** I'm already planning my costume. Think glitter and neon. And maybe a cape. Because if you're going to run a marathon, you might as well do it with flair.

**Speaker 3:** Glitter and neon? I'm in. Let's make it legendary.