

1 “More Structured and Organized for Learning”: Note Taking 2 on a User-Generated Video Sharing Platform - Bilibili 3

4 ANONYMOUS AUTHOR(S)
5

6 Video sharing platforms have emerged as extremely popular social media not only for entertainment, but are
7 also increasingly used for learning knowledge and everyday skills. Recently, some video sharing platforms
8 even introduced note-taking features to further foster such video-based learning (VBL). While note taking has
9 been extensively studied on specialized VBL platforms, its practice in social media settings (i.e., generic video
10 sharing platforms) remains unknown. In this paper, we present a qualitative study with 15 participants who
11 have engaged in note-taking practice on Bilibili, a popular social media platform featuring user-generated
12 videos in China and has recently launched a feature to support note-taking called BNNote. The study reveals
13 how the note is taken to structure, supplement, and substitute the video and how note-taking is collectively
14 carried out by the community as a self-organized whole, supporting a more structured and organized form
15 of social learning around a user-generated video sharing platform. We end by discussing our findings and
16 providing design recommendations to better support VBL on social media platforms.

17 CCS Concepts: • Human-centered computing → Empirical studies in HCI.
18

19 Additional Key Words and Phrases: User-Generated Video Sharing, Note-Taking, Video-Based Learning, Social
Learning, Bilibili

20 **ACM Reference Format:**

21 Anonymous Author(s). 2023. “More Structured and Organized for Learning”: Note Taking on a User-Generated
22 Video Sharing Platform - Bilibili. In *Woodstock ’23: ACM Symposium on Neural Gaze Detection, June 03–05,*
23 *Woodstock, NY*. ACM, New York, NY, USA, 26 pages. <https://doi.org/10.1234/1234567.1234567>

24 1 INTRODUCTION

25 Nowadays, more and more people are drawn to online video sharing platforms such as YouTube
26 and Bilibili, not only for entertainment, but also for video-based learning (VBL) [7, 33, 77]. For
27 instance, YouTube, as the world’s second most-used social platform with over 122 million daily
28 active users watching one billion hours [61], has attracted hundreds of millions of individuals
29 worldwide to enrich their online VBL activities [40, 64]. The 2022 annual report shows that 198
30 million people have been watching videos for learning some knowledge and skills on Bilibili [7], a
31 very popular video sharing platform in China, indicating that it has become a VBL environment
32 for many Chinese users, especially for young people [13].
33

34 To further support video-based learning, some video sharing platforms begin to enable users
35 to take notes on the platform. For instance, Bilibili has recently launched its note-taking feature
36 called BNNote (at the end of April 2021), allowing users to create, store, and even share notes while
37 watching videos without the need to switch screens [6]. In just several months, over 500,000 users
38 have posted more than 1.2 million notes on Bilibili. The new note-taking feature can potentially
39 help learners to more actively engage in learning activities [24, 32], since note-taking can trigger
40 learners’ active construction and processing of knowledge [24], help focus attention on the material

41 Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee
42 provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and
43 the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored.
44 Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires
45 prior specific permission and/or a fee. Request permissions from permissions@acm.org.

46 *Woodstock ’23, June 03–05, 2023, Woodstock, NY*

47 © 2023 Association for Computing Machinery.

48 ACM ISBN 978-1-4503-XXXX-X/23/01...\$15.00

49 <https://doi.org/10.1234/1234567.1234567>

50 presented [31, 35], and is commonly regarded as effective in enhancing learning performance [60]
51 and assessing learning progress [37].

52 However, while note-taking has been extensively studied [12, 16, 17, 20, 25], how it works in
53 today's extremely popular user-generated video sharing environments is still under-explored. In
54 previous work, note-taking is mainly examined in more specialized learning contexts [20, 34, 59],
55 focusing on its traditional cognitive facilitation of learning [17, 20, 81], but the video sharing
56 contexts are quite distinct. First, video sharing platforms feature non-professional users to freely
57 create, upload, and share videos for various purposes [22], not particularly for learning. This kind
58 of platform creates videos with broader topics but limited quality control [41, 48, 91], quite different
59 from platforms featuring professionally generated videos dedicated for learning, such as massive
60 open online courses (MOOCs). Second, unlike in traditional specialized learning environments
61 (e.g., classroom lecturing and online education) where instructors/institution provide support,
62 direction, and standards necessary for successful learning outcomes [14, 93], users on these video
63 sharing platforms take full responsibility for their own learning activities [42, 77]. With a range of
64 socialization and entertainment options provided on these platforms, learners can easily become
65 distracted with diverse video choices [64, 72], and not so well disciplined due to little external
66 support [9, 94].

67 To better understand note-taking in such a new context, we perform a qualitative study on
68 Bilibili. With its social media nature, popularity in considerable user bases, increasing use for
69 learning purposes, as well as recently released note-taking feature, it provides an ideal site for us
70 to examine note-taking related practice in today's user-generated video sharing based social media
71 context—outside traditional, more formalized, and specialized VBL setting. We report findings from
72 semi-structured interviews with 15 Bilibili participants who have posted notes on the platform. The
73 study sheds light on how note-taking is leveraged to support a more structured and organized form
74 of social learning on user-generated video sharing platforms, mainly by structuring, supplementing,
75 and substituting the video as well as collectively taking notes as a self-organized community whole.

76 Our work yields three contributions to the CSCW community:

- 77 • a first study of note-taking practice on today's extremely popular user-generated video
78 sharing based social media platform, beyond traditional, more formalized, and specialized
79 VBL settings.
- 80 • new empirical findings of the roles note taking plays for VBL on user-generated video
81 sharing platforms beyond traditional cognitive benefits. Specifically, we reveal that note-
82 taking is drawn on to support a more structured and organized form of social learning in
83 such a context, e.g., surfacing a layer of structure for user-generated videos, and scaffolding
84 a social force to stimulate more disciplined and organized learning for video-based learners.
- 85 • new design implications of how to better support social-oriented VBL in video sharing
86 contexts, in terms of communal ways of motivation, co-creating notes, and video and note
87 combined benefits.

88 2 RELATED WORK

90 2.1 Video-Based and Social Learning

92 Video-based learning, by definition, refers that people conduct learning and thus acquire knowledge
93 or skills via online videos [65]. Compared with typical classroom lecturing, it has multiple benefits.
94 Videos can serve learners by visualizing how something works [19], and display information and
95 details dynamically that are difficult to interpret through static text or photos [53], resulting in
96 improved learning outcomes [90]. Moreover, video can attract the attention of learners and inspire
97 them to communicate information around its content [16].

99 Various means have been explored to further enhance the VBL experience, such as video annotation and navigation. Video annotation is studied as an additional note of the video and a
100 simple way to provide feedback, promote reflection, and provoke discussion about video contents without modifying the resource itself [36, 54, 63, 87]. It can also be used to add extra explanations,
101 highlight points of interest, and pose questions for watchers [39]. In addition, video navigation techniques are also explored to better support viewer's interactions related to video segments, based on viewer annotation [16, 27], video transcripts [55], and viewer-video interaction data [38].
102 Annotation-based navigation, in particular, is based on the timestamps coming with annotations [27, 57], and is a two-way navigation method: once clicking an annotation, viewers can see that
103 video is jumping to its start time, and while the video player passes between its start and end time of the annotation, the annotation keeps activated and highlighted [16]. As such, it can help viewers obtain a summary of desired orientation, find places where certain concepts are explained, and review specific parts of the video they already viewed [55].
104

105 Due to the broad availability of online videos, VBL has become ubiquitous in specialized learning environments such as MOOCs [14, 28, 93], reinforcing and extending traditional teacher-led classroom experience [88]. In such learning settings, learners can access quality material at low cost and control where, what, and how to learn [83]. Meanwhile, they can receive necessary support from instructors or institution, such as certificate incentives [93]. As consumers of these professional VBL platforms, the majority of people perform individual learning [44]. Although with external support provided by instructors/institution, learners often feel isolated and disconnected due to a lack of social interaction from peers [28, 49].
106

107 By comparison, on social media sites including video sharing platforms, users are no longer merely viewers, recipients, or consumers, instead, they can actively contribute through connecting, communicating, and cooperating with one another and deriving value from this [75]. For example, video sharing platforms can scaffold social interactions around learning, such as time-anchored comments [82, 84], and learners on these platforms can also present their own and perceive others' learning traces via sharing [24, 25]. In such a context, learners can get exposed to each other's work and identify which ones are qualified as useful resources for learning [89], and thus improve their knowledge acquisition processes [3, 76]. Integrated with social media platforms, it can even avoid the potential risk of social isolation often associated with online learning [25, 71]. As such, video sharing platform provides a rich social learning environment, which enables learners to share outcomes (e.g., digital notes) and even collaborate with others while engaging in their own learning practice. However, as the video sharing platform is not designed specialized for learning, users can become easily distracted and disengaged with diversified videos and little external support such as instructors.
108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148 tools are available in the market, like Evernote¹, OneNote², GoodNotes³, and ColorNote⁴ for daily
149 use on the computer, iPad, and mobile contexts.

150 Diverse digital note-taking practice in modality, technology, and skill aspects has been studied in
151 previous research. For instance, Lin et al. study the distinction between formal and informal notes
152 [45], and indicate that formal notes support retrospective recall by reminding us what already
153 happened, while users commonly use informal notes as temporary storage for notable information
154 and prospective memory aids, focusing on current information and its future use. Chen et al. [17]
155 compare audio and text note-taking behaviors in mobile settings, finding that users take audio
156 notes with personal language to capture unfamiliar and emotion-related contents, while text notes
157 supply users with better accuracy and deeper reflection experience via modification. Bauer and
158 Koedinger [5] assess how different note-taking functionalities, i.e., typing, copy-paste, restricted
159 copy-paste, and menu-selection, affect student behavior and learning, noting a trade-off between
160 interventions to foster learning and interactions that note-takers really like. Zheng et al. examine
161 the note's core components as well as its dimensions of design space, and discuss design suggestions
162 to simplify repetitive note-taking tasks, furnish flexible generating processes, and lower entry
163 barriers [92]. Other scholars explore interfering strategies to develop learners' note-taking skills.
164 For example, Thaleia et al. verify the effectiveness of designing a game, which furnishes a legible
165 and entertaining learning atmosphere to cultivate note-taking skills for students in academic and
166 professional settings [21].

167 Collaborative note taking and sharing have also been extensively studied as a way to support
168 learners to exchange their thoughts and ideas with peers, and contribute to social learning [74].
169 For example, Fang et al. [25] design a digital note-taking and note-sharing tool, NoteCoStruct,
170 for online learners in VBL context, showing that note sharing can foster a sense of community
171 by providing learning traces shown to members. Moreover, collaborative note-taking is also an
172 effective method to stimulate social learning. For example, Kam et al. present Livenotes [34], a
173 cooperative learning system that supports group members to mutually interact by taking notes
174 cooperatively during video lectures, discovering that cooperative note-taking students collectively
175 create a more complete note and attend higher degree of group conversation than individual
176 note-taking ones. Collaborative note-taking can also enable learners to make more accurate and
177 comprehensive description of video lectures [26]. As such, collaborative note taking and sharing
178 contribute to better learner engagement, collaborative learning, and knowledge building [67].

179 At the same time, users' attitudes and practices of sharing notes with others have been reported
180 in the existing literature. Willett et al. report that participants are reluctant to share notes in the raw
181 form, because of feeling uneasy about their notes' structure and writing style and worrying about
182 forcing others to potentially read overloaded or irrelevant contents [30, 79]. As a result, people
183 probably perform pruning, reformatting, and editing works before sharing their notes or take notes
184 more prudently if they know them to be shared [79], some of whom even give up personal tones to
185 increase their notes' readability [51].

186 While digital notes have been extensively studied and the trend towards sharing notes in online
187 VBL contexts can be identified, however, most of the existing work is conducted in specialized
188 learning settings such as online education oriented platforms, and primarily concerned with
189 cognitive facilitation of note taking on user's understanding, reflection, and recall of videos, not
190 on broader meanings of note taking for VBL. In a word, how note taking works in today's video

191
192 ¹<http://www.evernote.com>

193 ²<http://office.microsoft.com/onenote>

194 ³<https://www.goodnotes.com/zh-cn>

195 ⁴<https://www.colornote.com>

sharing based social media context is still under-explored. In this paper, we address this gap by exploring the broader roles note taking plays for VBL in such a new context.

3 BILIBILI AND BNNOTE



Fig. 1. The BNNote list interface (right): (A) the “Notes” button, (B) the users’ note list, (C) the “Take a Note” button

Our study is based on Bilibili, one of the most popular online video sharing platforms in China—an average daily active user reaches 90.3 million [7], with nearly 80% of users aged under 24 [8]. Unlike specialized learning oriented video platforms such as MOOCs, Bilibili supports both learning and entertainment. In contrast to other popular video sharing platforms such as Douyin (a Chinese TikTok), whose feature is short-form videos lasting 15 seconds to a few minutes, Bilibili supports longer videos with even hours. Moreover, unlike YouTube, which was originally launched as a UGC (User Generated Content) platform, Bilibili has a history of transitioning from initially uploading ACG (Animation, Comics, and Games) to a UGC platform nowadays. Furthermore, Bilibili is a local UGC platform, whose video creators and viewers are mostly Chinese users with the same culture. With distinct social features such as Danmu⁵ and Virtual Coin⁶, as well as terms, rules, and symbolic language embedded in the video contents and comments, Bilibili has developed its own distinctive social culture in the local Chinese environment.

BNNote, refers to as the note-taking feature on Bilibili, was officially deployed in April 2021 [6] and conceived with the rise of pan-knowledge⁷. Since June 2020, Bilibili has launched the first-level “knowledge area”, which is integrated and upgraded from the original science and technology area. It is dominated by sharing knowledge, experience, skills, viewpoints, and humanities. In a short time, pan-knowledge content has jumped to the new windfall of this platform. Chen Rui, chairman and CEO of Bilibili, released a report at the 9th China Internet Audio & Video Convention

⁵A video commentary system that allows viewers to directly post comments overlaid on top of the video, flying on the screen as synchronously written texts with the video timeline

⁶A kind of virtual currency donated by viewers to the video uploaders for thanking their contribution

⁷Any forms of information that imparts a feeling of having learned knowledge or skills, without being restricted to a certain category



Fig. 2. The BNNote taking interface (right): (A) the screenshot function (with one click, the video frame at current moment of operation is captured, as shown in the note example), (B) the time-tag function (with one click, the time moment of current video playing is captured as shown in the note example, and then users can directly jump to the corresponding video frame through this time-tag), (C) the “Save” button, (D) the “Public Release” button, and (E) the “My Notes” button

held in June 2020, indicating that pan-knowledge content has accounted for 45% of the platform’s total broadcast volume [18]. The pan-knowledge was defined by him in this way: “As long as it improves personal skills and contributes to everyone’s life, it can be considered as pan-knowledge”. At present, the pan-knowledge content of Bilibili has spanned eight categories: Popular Science, Social Science & Law & Psychology, Humanities & History, Finance & Business, Campus Learning, Workplace, Design & Creativity, and Wild Skills. As such, the increasing focus on learning and pan-knowledge makes Bilibili ideal for us to understand video-based learning practice.

BNote was initially positioned to target pan-knowledge types of UGC videos. As a part of the video ecology, BNNote is designed to serve as an auxiliary tool for video watching and learning. Like BNNote, certain video comments can also assist viewers to understand key points of a video by summarizing its content [56]. Nevertheless, video comments are applied to broadly discuss knowledge, opinions, and thoughts with others in common form of short texts⁸ [1], while BNNote is intended to be created and shared tightly around the video itself, such as summarizing, simplifying, or supplementing video contents through longer texts and images⁹. Besides, only a few of video comments are about a summary of the video’s content, and these comments are interwoven in all comments placed below the video, making them easily overlooked [56]. In contrast, BNNote can be more readily recognized by both video watchers and note creators as a separate component. Therefore, we specifically focus on BNNote as our research subject.

In terms of functionality, BNNote is quite easy to use, which follows one of essential guidelines for the design of social software: keep the interface and function as simple as possible [75]. Beyond

⁸The content limit of video comments on Bilibili ranges from 2 to 2,000 words

⁹The content limit of BNNote ranges from 200 to 20,000 words during our study and the minimum limit was decreased to 100 words afterwards

295

To supplement video content: P6 made the note for a video he created himself, saying that he was for sharing his experience behind the AMV (Animated Music Video) production.

307

To substitute video watching:
The note could exist as a material for the video and yet independent of it.
For example, even if its original video disappears, the note P11 wrote can still remain.

320

Fig. 3. Screenshots of three participants' notes shown in their respective personal spaces (left top: P6's note, left bottom: P11's note, right: P8's note, and some practices of them can be corresponded to related quotes in subsequent Section 5.2)

324

some basic functions commonly found in note-taking products on the market, BNote is enhanced with two unique features (seen in box A and B of Figure 2, respectively): screenshot (the camera-like button) and time-tag (the flag-like button) to create a strong binding with the video. Regarding these two functions (seen in the note's two examples of Figure 2), with one screenshot click, the video frame at current moment of operation will be captured, while clicking the time-tag button, the time moment of current video playing will be captured, and then users can directly jump to the corresponding video frame through this time-tag. It is also its original purpose to associate notes with videos in lightweight ways, rather than developing an elaborate note-taking feature, which differentiates BNote from existing commercial or design tools such as OneNote, VideoSticker [12], etc.

336

To use BNote, users just open a video on Bilibili, and a “Notes” button will appear at the bottom right corner of the video (seen in A box of Figure 1). After clicking it, we are able to view, if any, all of users’ notes about the video (seen in B box of Figure 1). To create a note, the user simply clicks the “Take a Note” button (seen in C box of Figure 1). The interface will then switch to the note-taking interface shown in Figure 2. The note content will be saved automatically three seconds after the user stop typing, and he/she can also manually click the “Save” button (seen in C box of Figure 2). The note’s current saving status can be seen timely at the top of the BNote editor. If the user wants

343

344 to share the note after finishing, he/she can do so by pressing the “Public Release” button (seen in
345 D box of Figure 2). By clicking “My Notes” button (seen in E box of Figure 2), the user can jump
346 to the personal note space to see all his/her historical notes for further review or refinement, in
347 which the shared notes are additionally listed. For shared notes, others can see them in the note list
348 and the comment area¹⁰ of the corresponding video [6], and then click and go to these note spaces
349 to view details. As illustrated in Figure 3, we take screenshots of three participants’ shared notes
350 displayed in their respective personal spaces.

351 By early December 2021, more than 1.2 million notes from over 500,000 users have been released
352 on Bilibili, with about 120 public notes per day. Amongst them, shared notes account for less than
353 5% of total ones, with nearly 60% coming from PC. To expand the publicity, Bilibili has launched
354 multiple note-related clock-in activities¹¹ to date, aimed to encourage users to create and share
355 more notes, thus fostering a healthy and vibrant ecosystem of notes for VBL on Bilibili.

356

357 4 METHOD

358 4.1 Participants

359 We conducted a qualitative study to better understand note-taking related practice on Bilibili. The
360 only screening criteria was that participants had released at least one note on Bilibili, considering
361 the emerging nature of BNNote feature during our study. We designed a recruitment flyer that
362 describes the purpose, procedure and qualification of the study, and sent it to qualified users
363 through Bilibili’s messaging system. First, we chose random qualified users to send interview
364 invitations, and recruited P1 and P2 this way. Then, we made deliberate efforts to recruit active
365 note users who could provide richer experiences, by contacting the top-30 note users ranked by
366 the number of comments, and recruited P3, P4, P7, and P8. P8 also recommended P5 who was the
367 product manager of Bilibili, and P5 further recommended P9 and P11 whose notes were related
368 to course learning. We also contacted 6 prize winners of Bilibili’s initial note clock-in activities
369 and recruited P12–P15 (P8 was also in the list of prize winners). We stopped when reaching data
370 saturation [50]. In the end, we recruited a total of 15 Bilibili users as our participants, with a varied
371 levels of engagement, including one short attempt merely as well as having shared multiple notes.
372

373 Participants’ demographic information is shown in Table 1. These 15 participants, 7 male and 8
374 female, aged 15–26 years (average age of 20.3), were distributed in nine different provinces around
375 China. The number of notes shared by participants and their total number of notes were based on
376 the data available at the time of our interview. As seen from Table 1, even P8, who was deemed
377 by Bilibili’s product department as one of the most active note-taking users, as well as P15, the
378 champion of note clock-in activity, only took notes in dozens of magnitudes. This could be because,
379 at the time of our study, the BNNote feature had been officially published and propagated for just a
380 few months and participants had not yet produced so many notes. Several participants have since
381 made and shared more notes. Unlike the other participants who took notes on the videos they
382 consumed, P6 made the note for a video he created himself. Although we tried our best to balance
383 the age and occupation distribution, the majority of participants were school students and young
384 people who had recently entered society, which was consistent with Bilibili’s user population,
385 mostly young people in their 20s [8].

386

387

388 ¹⁰In the comment area of the video, every note is specifically marked with the “Note” word and a note-taking icon, and
389 people can click “Open the Note” to see its details

390 ¹¹An open activity organized by Bilibili that requires involved users to write and publish notes limited to a minimum
391 number of days within the specified period, and rewards the winners who have made the highest-quality notes with prizes
or virtual currencies

392

Table 1. Demographics of Participants

ID	Gender	Age	Occupation	City (Province)	No. of Shared Notes (Total Notes)
P1	F	22	Postgraduate	Wuhan (Hubei)	2 (3)
P2	F	18	Senior School Student	Fuzhou (Fujian)	1 (17)
P3	F	21	Undergraduate	Unknown (Anhui)	2 (2)
P4	F	15	Senior School Student	Jinhua (Guangdong)	1 (3)
P5	M	24	Bilibili Product Manager (Master)	Shanghai	1 (1)
P6	M	26	Postgraduate	Zhengzhou (Henan)	1 (1)
P7	M	17	Senior School Student	Unknown (Anhui)	2 (2)
P8	M	21	Taobao Shopkeeper	Wuhan (Hubei)	24 (31)
P9	M	17	Senior School Student	Haerbin (Heilongjiang)	8 (8)
P10	F	22	Nurse (Bachelor)	Jiaozuo (Henan)	1 (1)
P11	F	17	Senior School Student	Jibao (Shanxi)	1 (1)
P12	F	20	Undergraduate	Wuhan (Hubei)	9 (18)
P13	F	23	Accountant (Bachelor)	Jiangmen (Guangdong)	13 (21)
P14	M	24	Undergraduate	Shenyang (Liaoning)	17 (19)
P15	M	17	Senior School Student	Shenyang (Liaoning)	15 (24)

4.2 Procedure

Due to dispersed distribution of our participants, we conducted semi-structured interviews primarily by voice calls, from late July to early October 2021. Our interview questions covered participants' experiences and details about their concrete note-taking cases, including how they noticed the BNNote feature, the context in which they took a note, the most impressive or latest note-taking experience, why they chose to share their notes, any concerns they have about sharing notes, users' note reading and learning experiences on Bilibili, and so on. When we discovered some intriguing or illustrative points, some participants also agreed to share specific screenshots with us. Despite using the same interview guideline for all participants, we usually refined our questions or the inquiring order based on their answers accordingly, to ensure a natural conversation flow and to collect data as rich as possible based on each individual's unique experience, until saturation was reached.

Each interview lasted around 60–90 minutes in Mandarin Chinese language, and each interviewer was compensated with 50 RMB for their contribution to the study. Our study was approved by the authors' institution. Prior to each interview, we clearly stated our intentions and offered information about the study. All data were audio-recorded and then transcribed with the permission of our participants, and were anonymized for analysis in this paper. Specially, under-age participants were interviewed at home with the permission and presence of their parents.

4.3 Data Analysis

We adopted a grounded theory approach for data analysis [69]. An open coding was conducted by authors on the transcript independently. We went through the transcript line by line, summarizing the points in the data using a higher level of code abstraction. Through the axial coding conducted with iterative discussions among authors, we attempted to find links between the codes, sorted and grouped them into sub-categories and further into categories. Finally, with the selective coding, we identified a few categories that represented the rest of sub-themes as recurrent and central themes. As the analytical process progresses, the roles of note taking plays in complementing drawbacks of user-generated video and scaffolding a social way of learning gradually emerged as a major theme. Representative quotes were translated into English from Mandarin and used as illustration in this paper, which we will explain further in the following section.

442 5 FINDINGS

443 In this section, we will summarize some background information about the participants' video-
444 based learning involved on Bilibili, and then elaborate on various roles note taking plays in the
445 context of video sharing platforms, including complementing the drawbacks of user-generated
446 video as well as supporting a social way of learning by collective conduct on the platform.
447

448 5.1 Overview

449 All our participants perceive Bilibili as a platform for informal learning, as it supports freely search,
450 transmit, and acquire knowledge or skills via the materials generated by non-professional users.
451 Videos on Bilibili are seen as for both learning and entertainment, as opposed to those on other
452 professional video platforms specialized for learning with instructor/institution involvement and
453 even curricular criteria presence, such as IMOOC¹² (an online training platform for IT skills) and
454 YuanFuDao¹³ (an online education platform that supports under-age students to learn various
455 school subjects).
456

457 From the note related data, we learn that adults watch a wide variety of videos depending on
458 their individual interests or skill improvement needs, while under-age participants predominately
459 watch school course videos. This phenomenon could be related to the particular social context
460 where study is highly emphasized for Chinese adolescent students, so social media is also drawn on
461 for learning outside the classroom. Moreover, participants who have made just one note perceive
462 BNNote as being quite similar to other Bilibili features that they have been using regularly, such as
463 video comments (P10) or columns (P6, P11), and we have learned that they continue using these
464 familiar features in addition to BNNote.
465

466 We find 9 of 15 participants are stimulated to write the first note by seeing others' notes, which
467 can be explained with social learning theory that individuals act influenced by observing others'
468 behaviors [3]. For example, P11 is motivated when seeing others' notes are well-detailed, showing
469 how much effort has been put into learning, so she decides to take notes too. P6 is influenced by
470 the uploaders he follows: "*I saw an uploader I follow was using BNNote to publish a note this year,*
471 *so I thought I could use it too.*" In addition, the aforementioned clock-in activity designed by the
472 platform to encourage note-taking also plays a role here. For instance, P15 has become fully aware
473 of the BNNote feature after seeing many notes left in the video's comment area, and finally decides
474 to take one himself with a timely push of the clock-in activity from the platform.
475

476 All participants perform varied degrees of editing works to make the note more readable before
477 sharing it to others, consistent with findings from previous work [51, 79]. Moreover, all participants
478 preserve their shared notes until now as a convenient means of marking, reminding, and guiding
479 what they have learned. Whereas, the private notes (i.e., these notes have not been shared by
480 users and can only be viewed by themselves) are characterized by fragmented, scribble-like, and
481 couple-of-word. As the temporary storage, the majority of private notes have been later deleted by
482 participants, aligning well with the lifecycle of formal and informal notes mentioned in existing
483 studies [45].
484

485 Now we will turn to elaborate how note taking is drawn on to complement the drawbacks for
486 video-based learning by structuring, supplementing, and substituting the video, as well as scaffold
487 a social way of learning by sharing notes with the community, collectively co-creating notes, and
488 mutually augmenting note-taking skills.
489

490 ¹²<https://www.imooc.com>

¹³<https://www.yuanfudao.com>

491 **5.2 To Complement User-Generated Video**

492 Videos contain rich information in audible and pictorial forms, and have become increasingly
493 popular resources for learning [12]. However, videos on social media are mostly generated by
494 ordinary users [22, 48], and not necessarily specialized for learning as well. So, the structure of
495 these videos is sometimes not so evident as professionally produced ones for learners, and it is time
496 consuming to watch and learn from the video [46, 66]. From our study, we found note taking is not
497 simply to cognitively facilitate learning, but many times, to compensate the various disadvantages
498 of learning on video sharing platforms, particularly in terms of structuring, supplementing, and
499 substituting the video in a cost-effective means.

500
501 **5.2.1 To Structure the Video.** In addition to its less obvious structure on social media aforementioned,
502 video is essentially a temporal stream of data containing both audio and pictures, making
503 it difficult to see the structure and quickly get an overview. As such, its structure, if any, is less
504 intuitive than other media forms such as pictures and texts that unfold in the visual space. P12
505 compares it with text this way:

506 *The video is a stream. I have to drag the progress bar if I want to go to view something.
507 It is extremely troublesome this way. It is very unlike the text, which is so clear at a
508 glance.*

509 With note taking using BNote, however, it enables users to add a layer of structure to the video
510 stream. First, it is through the two functions of BNote, i.e., time-tag and screenshot, which are
511 perceived by 11 of 15 participants as the most useful, as they can be used not only to externalize the
512 structure contained in the video but also to tightly correlate with particular video segments. For
513 instance, P13 described how she used the time-tag function in a note on medical popular science
514 regarding common cold:

515 *On a certain video frame, the speaker usually said what we would discuss next, and
516 then I clicked [a time-tag button to add it to my note] right away ...*

517 Similarly, adding the time-tag makes it easier to see different video segments, as P14 explained:

518 *This video (on cities and cultural heritage) is segmented. The first and second episodes,
519 for example, are about an introduction. Just like the previous video on grammatical
520 tenses, it (this video) involves several episodes, each of which tells different contents.
521 So, it was expressed more clearly by adding the time-tag [that corresponds to the video
522 segment] like this.*

523 Moreover, with the time-tag function of BNote, people can quickly locate certain points of the
524 video, making it more convenient and more intuitive for them to navigate, which is similar to the
525 timestamp in video annotation [27, 57]. Our participants commented that BNote made it easier not
526 only to see the structure but also to quickly jump to the part they wanted to see. P1 described this
527 way:

528 *It (time-tag) is equivalent to a quick index. For example, in the beauty makeup [video],
529 the first part is about the lipstick, and the second part is about the foundation. [At this
530 time], I can quickly go to where I desire with these time markers. That is, if I want to
531 learn more about a certain part, I can click [the time-tag] and find the corresponding
532 part that I want to see.*

533 At the same time, people also often use the function of screenshot to make the notes even more
534 illustrative and visually appealing. For example, besides the time-tag, P11 also used screenshots
535 along headlines to make them more eye-catching. Similarly, some screenshots serve as a quick
536 summary of the video's topics, as P9 mentioned:

540 *They mostly reflect some key contents of this video (about organelle). That is, the*
 541 *key points of this video are these [screenshots]. For example, [as shown in the first*
 542 *screenshot], the biological membrane system is divided into nuclear and cell membranes,*
 543 *which are both highly important. And then, the cell membrane can be divided again.*

544 P13 utilized screenshots to select representative examples from the video to illustrate specific
 545 significant topics. P11 described a case of how she used screenshot and time-tag functions together
 546 to mark things as future references for a video story (“Kuafu Chased the Sun”) she adored.

547 What is more, users also externalize certain core pieces of information in a textual form, such as
 548 topics, highlights, and more, so to make the video’s structure more visible, mainly through BNNote’s
 549 various textual features such as font color, background color, and font size. For example, P9 used
 550 font and background colors to distinguish different levels of content in the video:

551 *The red font is to determine whether it is vector or scalar, which is the focus of this*
 552 *video (about physical concepts). The blue font is to express an explanation of these*
 553 *concepts. The green font means a subheading ... [By doing so], you can tell what’s more*
 554 *important, right? ... The fonts with yellow backgrounds are key points, such as, “the*
 555 *timeline doesn’t have a negative half-axis” is a relatively important thing ... Yes, these*
 556 *fonts with red backgrounds are the most important.*

557 P1, on the other hand, leveraged boldness to highlight what is important in the video on popular
 558 science about diarrhea:

559 *These (bold fonts) are the [video’s] knowledge highlights, which allow me to rapidly*
 560 *recognize that “diarrhea” belongs to “visceral abdominal pain”. Further, what is “visceral*
 561 *abdominal pain”? That is, it’s “stomach + intestine”, and something in the intestine*
 562 *triggers the pain receptors, that’s why I feel pain. By doing so, I can see the focus at a*
 563 *glance ... Others (unmarked contents) are secondary knowledge that can be considered*
 564 *as common sense, and hence there is no need to mark them.*

565 P12 also detailed how font-color or font-size of BNNote can be used to distinguish video’s concepts,
 566 e.g. using cyan fonts as glossary for some special or important concepts, and small gray fonts for
 567 supplement and explanation of those concepts.

568 As mentioned above, users were able to surface the video structure by using a variety of BNNote
 569 techniques. That is, the textual note, combined with BNNote’s functions such as screenshot and
 570 time-tag, provides an easy and visually appealing way for people to see the structure and navigate
 571 the video. Although video annotations can also structure and navigate the video to some extent
 572 by highlighting key points [39, 55] and timestamp-based navigation [27, 57], they are usually
 573 incomplete, scattered, and intertwined with the video.

574 **5.2.2 To Supplement Video Content.** Since the user-generated characteristics of videos on social
 575 media, the content of video may not be as clarified as professional oriented ones. In such a con-
 576 text, not merely to transcribe what is said from a video, the note is also commonly drawn on to
 577 complement the video stream to make it more organized and enriched. We discovered that people
 578 endeavored to refine, replenish, and expand contents that video lacks, through a variety of source
 579 avenues, supplementary methods, and screening strategies.

580 Very often, participants attempted to supplement the video content with more detailed textual
 581 descriptions by BNNote’s textual functions, as shown in P10’s case of cooking note:

582 *In preparing dishes, for example, the speaker didn’t say how to cut onions. She didn’t*
 583 *say anything and just started cutting. So, I wrote how to cut it in the note. Also, the*
 584 *whole process of frying vegetables was not fully described, [or] she might describe a*
 585 *part of it. Anyway, I followed the steps in the video and added a text description to it.*

589 P9 was taking some physics lessons from Bilibili for preparing new knowledge involved in the
590 upcoming semester, and sometimes he provided examples of his own by utilizing BNNote's image
591 uploading function if the video lecturer did not provide one or the ones used by the lecturer were
592 not so understandable.

593 As for the complementary materials, nearly half of participants reported that they commonly
594 leveraged relevant comments from the video comment section as a source. P8 explained why he
595 included two creatures not mentioned in a biological popular science video to his note:

596 *Because someone played a gag with these two creatures in the comment area, so I added
597 these two. That is, the Olympic champion claimed that XiaoLiang (the video uploader)
598 is a "Tibetan Rat" at the beginning of the video, and well, then a lot of people started
599 saying "Tibetan Rat", "Tibetan Rat", ..., so I added "Tibetan hamster" and "Moupin
600 Pika", which weren't mentioned in the video.*

601 P11 also detailed how she gleaned comments to list reasons for objecting to a particular viewpoint
602 in her note about "The Peach Colony", which is a very famous writing in classical Chinese:
603

604 *The No. 1 (reason why Utopian ghost stories are untrustworthy) is that I saw a hot
605 comment, saying he/she had read the "The Peach Colony", "Doraemon", and other
606 works, and believed these very good classics had been maliciously interpreted. So, for
607 this article [explained in this video], malicious interpretation seems very relevant to
608 list ...*

609 P2 used the video uploader's profile as well as representative comments as a source to complement
610 the tutorial video about slide design:

611 *I have referred to the representative comments in the comment area below this video, as
612 there may have some summaries, ... Also, there is some information in the uploader's
613 profile, which I think is very valuable to understand what is delivered in the video, so
614 they can't be ignored.*

615 In one medical popular science video, P15 noticed a comment with references provided by the video
616 uploader himself, so added the references to his note too. When incorporating comments, some
617 screening and collecting strategies were employed. For example, P2 usually filtered out those with
618 advertising feelings, or those simply for fun. P12 exhaustively collected the comments of one class
619 representative¹⁴ who had provided all of the video summaries but distributed in scattered spots.
620

621 In addition to video comments, participants also often located pertinent information online, and
622 complemented it with the video. For example, P8 added information from the Internet that wasn't
623 covered detailedly in the video:

624 *The speaker didn't mention features and so forth of a creature (in this biological popular
625 science video), so I went to Baidu to find relevant contents. For example, for the second
626 creature ("indigo fruit"), the speaker mainly discussed what it is, [but] he didn't talk
627 in detail, so I looked up its related features on the Internet ... The video explained the
628 reasons and principles of "Kangaroo bending waist", and the introduction and origin
629 [of Kangaroo] were what I found in the Baidu encyclopedia.*

630 P15 added correlative contents from searching the articles of WeChat Official Accounts in his note
631 on how to solve insomnia problems. In the note about cities and cultural heritage, P14 provided
632 pertinent materials from his school's website that correspond to the video's theme, such as content
633 about the history and growth of the city called Panlongcheng. P11 reported a case of how he
634 combined Danmu and search engine to add materials:

635 ¹⁴Someone who is recognized by others as a video's class representative, because he/she summarizes the content of the
636 video well, in the form of comment or note

638 *On Danmu, someone recommended [us] to see the “Peach Blossom Source Poems”, so*
639 *I went to find the hyperlink and posted it to my note ... The Baidu encyclopedia and*
640 *links to this ancient poetry were added at the end ...*

641 Users even used BNNote to share their personal experiences and thoughts to complement the
642 video. For example, P6 is an anime fan and utilized BNNote to share his experience behind the AMV
643 (Animated Music Video) production he shared on Bilibili:

644 *In order to let the audience better understand some ideas behind my video creation, or*
645 *what I wanted to express ... I wrote this note mainly for sharing my own experience,*
646 *and my understanding of animation.*

648 People also added their personal feelings and insights to their notes, such as appreciation for the
649 video’s speaker, desires for encouragement and discussion, and so on. For example, P12 stated so at
650 the start of her note on “Aesthetic Principle”:

651 *The speaker explains the course in a simple and understandable way, and I feel quite*
652 *rewarded. If you like my note, you can collect it from my column. If you have any*
653 *problem, I will be excited to discuss it together.*

654 To sum up, people tried to fill in any content gaps in the video as much as possible by improving,
655 replenishing, and expanding methods. Not only did they use readily available and accessible
656 information such as video comments, but they also proactively sought and added multiple materials
657 via any online avenue they could think of. Besides, certain participants even added their personal
658 opinions about the video. They built their own strategies for filtering and gathering additional
659 information during the process.

661 5.2.3 *To Substitute Video Watching.* Compared with a video stream, the note is taken as a media
662 that can convey knowledge in a more concise form, which at times provides as a more selective
663 way of viewing by substituting video watching. P10 commented that a note combining screenshots
664 and words was very intuitive to view:

665 *It looks quite intuitive. For example, to prepare the ingredients [for cooking], this picture*
666 *can exactly depict the container containing ingredients as well as the spoon. [In this*
667 *way], with picture and text together, it looks so intuitive ... Just like I said, how big*
668 *is this spoon? It’s something that could be difficult to explain in words. For example,*
669 *I can’t say it’s a round spoon with a diameter of 3~5 cm, right? With the screenshot,*
670 *[together with the text], I see the information intuitively.*

671 P12 said, instead of watching the video, she could directly read pivotal information from the textual
672 note, since “*It (the note) is equivalent to translating the video to text, representing the knowledge in a*
673 *more condensed form.*” P7 mentioned that the note was especially valuable and beneficial for those
674 who couldn’t watch the video or when it was just not so convenient for them to view. As such,
675 with a mixture of picture and text in a static, condensed, and spatial form, the note can be often
676 utilized as a more selective way to replace video watching.

677 When review is concerned, while the video also works as a carrier of knowledge, reviewing
678 notes is definitely considered more efficient than re-watching the video again. When participants
679 were asked how they reviewed the materials related to the video, they answered that they often
680 just reviewed the note directly instead of re-watching the video. P2 shared her experience:

682 *Almost always, when I saw something and thought it was useful, I used the screenshot*
683 *function of BNNote to put it inside the notes. This way, if I need something [relevant],*
684 *I can just look at my notes [that contains screenshots] without having to watch that*
685 *video again.*

687 P9 reported that he went over his notes several times, instead of watching the video repeatedly to
688 memorize something:

689 *I reviewed it (the note) several times and then remembered some important points. I
690 mean, I memorized it again and got the important parts of the note into my head ...
691 instead of watching the video again and again.*

692 While the cognitive effects of notes for knowledge review has been commonly reported [4], in
693 the context of video sharing platforms, we also found that people commonly drew on the note as a
694 shortcut to preview the video so help them decide whether to watch the entire video or not. As
695 videos generated by users on social media are highly diversified [64, 72], it helps learners make
696 efficient video choices on such a platform this way. For example, P13 reported that she usually read
697 notes to decide whether or how to watch the video:

698 *It may take some time to watch a video, but if you can see the notes, it will be clear
699 which part you want to watch the most, and you can then focus on that part only [by
700 clicking the corresponding time-tag], so you can spend less time to understand what the
701 video is about ... I'm too busy, my time is very valuable, because I work in a hospital ...*

702 Meanwhile, P13 mentioned that she usually browsed the overall structure of other's notes, and
703 then decided to view the video if there were any points that interest her.

704 Our participants also liked that the note could exist as a material for the video and yet independent
705 of it, which is different from video annotation [39, 57]. P3 said: “One advantage of BNNote is that
706 even if its original video disappears, the note I wrote can still remain.” BNNote also works well as a
707 memory aid [45] that is convenient to carry when people need to put the knowledge into practice.
708 For example, P4 was practicing singing and dancing. The note with lyrics was helpful for her to
709 practice singing, and movement screenshots helped her to practice dancing:

710 *I wrote down [the lyrics of this song] because it's convenient for me to look at the lyrics
711 and practice singing it out on my own ... I learn dancing on Bilibili, and when I came
712 across the repetitive movement like this, I saved it as a screenshot [and thus used it] to
713 follow up the movements.*

714 Similarly, P10 is interested in cooking, and relied on the note for aid whenever she wanted to cook,
715 instead of watching the video again.

716 From the above cases, we can see that in many situations, including knowledge review, practicing
717 knowledge from the video, and making watching decisions, participants often leverage the “easy to
718 glance” and portability [11] nature of the note to replace actual viewing. In this way, although both
719 video and note can work as knowledge carrier, the note, as a lightweight alternative carrier of video,
720 is often preferred, which provides as a more selective way of viewing and learning knowledge on
721 social media.

722 **5.3 To Scaffold a Social Way of Learning**

723 Bilibili is essentially an online community [47], where members come to watch and learn videos
724 together [13, 29, 86]. That is, video watching and learning are not in isolated forms but embedded
725 within the social process of the community. Our study also reveals the social function of note taking,
726 in which participants' note-taking practices are similarly shaped by the social process in multiple
727 ways. Specifically, note-taking scaffolds a social way for Bilibili users to get more disciplined, more
728 organized, and more high-quality produced for learning, by sharing notes with the community,
729 collectively co-creating notes, and mutually augmenting note-taking skills.

730 **5.3.1 More Motivated and Disciplined through Shared Note Taking.** On such social platforms not
731 specialized for learning as Bilibili, it is especially important for learners to leverage the power
732

of members within the community to facilitate video-based learning, as there lacks supervisory support from outsiders (e.g., instructors). Here, we notice that shared note taking enables them to leverage such social force for stimulating learning. Specifically, after motivated by others' notes to take and share their first note, our participants had it in mind that their notes were also to be shared and seen by others, which then gave them an extra stimuli to be more disciplined and accountable [79]. For example, P1 made a conscious effort to use the public audience as an external force to push herself to ensure the quality of her own notes and be more focused on learning Go language programming:

I hoped I could be more focused when watching the video for learning ... I felt it (using BNNote to take notes for the public) gave me a sense of being monitored, which made me take it more seriously. During the note-taking process, as I had envisaged making it public, I held more responsibility for these notes, and I couldn't just scribble anything. In a sense, [public] note-taking served as a function of monitoring, with which I was forced to take it more seriously during the whole process.

Likewise, P7 saw others' positive comments as an affirmation of his hard work, which then punched him to keep studying English during the summer break:

Yeah, this is my [greatest] motivation to continue writing this note and learning [English words], ... When I saw someone find it very useful, or someone urged me when to update my note, I felt quite worthwhile. They [the comments] let me feel so motivated to memorize [and write down these] English words every day.

The clock-in activity of note-taking organized by the platform was further leveraged to help them more engaged in learning. For instance, P13 described how participating in the note-taking activity helped her stay focused and engaged in studying drawing, which she had always meant to do but was not so successful:

I did it not only to win this clock-in prize, but also to get myself moving, learning, and finding my life goals ... Yes, it's true, I was so lost after graduation that I didn't know what to do. I wanted to study, but I wasn't motivated enough. So I participated in the clock-in activity to keep myself engaged ... It did get me started and making progress ... Right! I've been relying on the activity to force myself to watch tutorials and learn drawing, and I haven't watched any entertainment videos since then.

As shown here, BNNote, together with the clock-in activity organized by the platform were actively utilized by these learners so to keep themselves more focused and more engaged in VBL on Bilibili by taking quality and regular notes.

5.3.2 *Self-organization in Collective Note Taking.* We found that it was also a collective, not merely individual, activity for note taking using BNNote, in that people worked together to make the shared notes more complete and better, as similarly found in previous work [34, 67]. In particular, if they noticed that someone had taken a note for a video, and if the note was an excellent one, they would just make use of it instead of creating a new one. On the other hand, if they found that a video was worth note-taking and yet no note was taken or something was missing in existing notes, they would spontaneously supplement it by creating a new note or adding new information to their own. From the study, it appears, these are the tacit rules people follow, and it is these rules that prompt people to spontaneously and collectively complete and improve notes in a more organized manner. P10, who often watched popular science videos, described such a situation:

For example, I watched these types of popular science videos, such as those uploaded by XiaoLiang (a famous biological science uploader) and TuBaGe (a famous medical science uploader). I might watch them after the video was probably uploaded two or

785 *three days. In this case, someone else usually had taken notes, and it was nice for me to*
786 *simply view or collect these notes. I didn't have to do it myself anymore in this case.*

787 P11 similarly said that if there were detailed notes already, she would just read their notes instead
788 of creating new ones.

789 So it is a common practice that before taking notes themselves, they check whether there are
790 any existing notes, or how complete the notes are if there have any. P11 even checked whether
791 there were any good comments too before she started to do it herself:

792 *When I saw this video (on “The Peach Colony”), I first turned to the comment section*
793 *to see if there were any high-quality comments, if anyone had summarized it, and if*
794 *there were any representative comments. After seeing no one did this, then I clicked the*
795 *note list, and found no one had made any notes there either ... After a quick browse, I*
796 *thought the video was logically clear. So, I just started to take notes right away ...*

797 In one of her notes on literary theory, P12 reported that she added some content that she believed
798 was important but had not yet been included in others' notes:

800 *In the first half of the video, the speaker talks about some cutting-edge content [before*
801 *starting with the introduction], but I found that no one had written it down. Basically,*
802 *others started taking notes directly from the introduction, but I thought it could be*
803 *recorded as a course beginning.*

804 When seeing others' concise notes with only a outline, P12 decided to supplement it in her note
805 with more contents:

806 *Take the note of “Aesthetic Principle” as an example, if we didn't understand what the*
807 *video speaker explained beforehand, it would be difficult to read such notes. So, based*
808 *on those, I supplemented mine ... That is, I added crucial contents, examples, and logical*
809 *transitions of the video based on the outline.*

810 To many, using others' notes and making notes for others are a form of mutual help in the
811 community. P9 explicitly put it that as he derived value from others' notes, he would also make
812 notes for others:

813 *I started watching the video (about physical knowledge) from the first chapter, but only*
814 *started making the notes later, as someone else had made some notes for the first few*
815 *chapters. So, I didn't need to repeat it ... Sometimes, I'm a little lazy. If there have been*
816 *notes already, I needn't take the note and can just use their notes to memorize things.*
817 *At the same time, the notes I took could also help others in this way.*

818 P3 used the textual feature of BNote to transcribe songs to share the lyrics mainly for her singing
819 friend:

820 *It was mainly because I knew a singing uploader friend, and I was actually helping*
821 *her to transcribe this song so she didn't need to do it [and could just cover this song by*
822 *reading my note]. That was just a help between friends.*

823 5.3.3 *Mutual Learning for Note Taking.* By using others' notes and sharing their own ones, it has
824 also become a social process for users to learn from each other about note taking. This can be
825 viewed as a more cost-effective way for cultivating note-taking skills, compared with those through
826 existing gamification techniques [21, 62]. More specifically, they paid attention to others' notes,
827 compared the difference, e.g. in quality, and thus learned and internalized the note-taking skills
828 from others to strengthen their own notes to be more structured and organized.

829 When asked why they viewed others' notes, 12 of 15 participants mentioned that it was for
830 learning—learning note-taking skills besides learning the video materials. P14 described it this way:

834 *I learned others' note-taking skills, including what content was marked down, what
 835 content was added, or other similar skills like this. I'd like to see how they took notes,
 836 or the skills they used to make notes more understandable, and then tried to master
 837 these skills myself.*

838 P11 tried to compare notes to improve her own notes:

839 *If others had taken notes, I usually looked at theirs to see what differences there were
 840 between their notes and mine, whether there were any points to learn, and what could
 841 be used to improve it.*

843 P9 also drew on others' notes to see whether there were anything missing such as structure and
 844 organization in his own notes, and gradually improved these skills so the notes he took could be
 845 more acceptable to others:

846 *I learned from them (others' notes on the knowledge of different subjects) to improve
 847 my own notes, such as the focus, labeling, and layout. By doing so, I can gradually
 848 align myself with the public interest.*

849 For picking up note-taking skills, people used different strategies. Some viewed others' entire
 850 notes including structure, focus, and details, such as what P13 did: "First, I viewed the overall content.
 851 Then, I selected the contents that arouse my interest to view the details." P1, on the other hand, assumed
 852 that all she needed to do was locate the content she desired:

854 *I only needed to take in the information that was beneficial to me [and added it in my
 855 note], right? I wouldn't pay attention to any extra information.*

856 Others made specific efforts to find high-quality notes to learn, e.g. how to organize their own
 857 ones better. For example, P8 drew on people's feedback such as the number of likes to identify
 858 high-quality notes:

860 *I'd go to see who has recently released notes based on the number of Likes. I'd see how
 861 well others wrote ... For example, I viewed what forms others used, such as how the title
 862 was written.*

863 P13 told us the reason she marked time segments that emerge video knowledge points at the
 864 beginning in a note on common cold to make the note's structure more obvious, was because it
 865 was used by a popular note taker (it was P8 actually), also called "class representative" on Bilibili:

866 *Because I've used others' notes for reference, and saw one note was taken like that, and
 867 the speaker listed time tags for different video segments at the beginning ... It was from
 868 a "class representative", who seems to have written a lot of good notes, and took them
 869 special seriously ...*

871 P2 expressed her desire to obtain more high-quality notes for reference, but did not find many of
 872 them. Similarly, 9 participants regarded seeing high-quality notes as an effective way to enrich
 873 their note-taking abilities.

874 Taken together, note-taking has been collaboratively and spontaneously carried out by the
 875 community of members. Throughout the process, it provides a social way of getting more disciplined,
 876 more organized, and more quality produced for these members' learning on Bilibili.

877 6 DISCUSSIONS AND IMPLICATIONS

879 In the preceding section, we identified the various roles that note taking plays on Bilibili, specially
 880 in its surfacing the structure of user-generated videos and scaffolding a more organized social way
 881 of learning. Below, we will discuss our key findings and implications.

883 6.1 The New Roles of Note Taking

884 While note taking, and digital note taking in particular, has been extensively examined, it is mainly
885 explored in a more traditional, formalized, and specialized learning context. In this paper, however,
886 by focusing on video sharing based social media platforms, we uncovered new roles note taking
887 plays for VBL. As we can see in our study, in this particular context, note-taking is in many
888 ways drawn on to alleviate the shortages of VBL on the platform as well as take the advantage
889 of the platform to enhance the social ways of learning, e.g., making the less structured videos
890 more structured, and leveraging the social forces for more disciplined and organized self-driving
891 learning:

892 While video annotation can also structure the video to some extent by highlighting key points
893 [16, 38, 54], however, as shown in the study, compared to the annotation which is tightly intertwined
894 with the video, the note, although is created from the video, often works as a new user creation
895 that can exist independently of the video. Compared to annotations which are often short or even
896 scattered, the note typically contains long and complete contents combined with texts and pictures.
897 This form of independent existence is valuable too for VBL on social media platforms, e.g., as a
898 preview to the video, as a form that is easy to carry and review, and as an alternative for video
899 watching, etc. Here, both videos and notes serve as recordings of knowledge, so the roles of note
900 taking is not merely about recording or storage as one main need as in traditional learning settings
901 [23, 37]. As revealed in our study, it is the different characteristics inherent in video and text/image
902 that matter here in shaping how people make use of them for learning purposes. As shown in
903 the study, video and text/image have different strengths and weaknesses, and people use notes
904 to draw on their strengths and weaknesses to make them complement each other. For example,
905 participants leverage the note's spatial visual form to reveal the structure of video, exploit the
906 time-tag feature to navigate the video, and utilize the structural note for preview or review before
907 engaging with the video to provide a more selective way of viewing. As such, what BNNote provides
908 is more flexibility to use these two media forms to fit different situations, as a facilitator or substitute
909 for video watching and learning.

910 Moreover, note-taking is also leveraged to turn Bilibili, a social media platform, into a learning
911 environment by cultivating a social atmosphere for learning. Unlike specialized VBL contexts
912 such as MOOCs [14, 93], video sharing platforms such as Bilibili were not originally designed for
913 learning, but mainly as social media for entertainment and socialization. To conduct learning in
914 such a social media context, especially social learning as uncovered in our study, it poses more
915 challenges for users to keep focused and disciplined, e.g., the use of social media can be easily
916 distracting. However, as shown in the study, by viewing others' notes and sharing their own notes,
917 our participants make the learning activity more visible, and invite extra social forces to keep them
918 more motivated and disciplined.

921 6.2 Note-Taking as a Self-Organized Whole on Bilibili

922 Our findings illuminate the dynamic whole emerged through people's self-organizing behaviors on
923 Bilibili, not imposed by some external forces such as instructors, which, in many ways, resonates
924 with the Gestalt theory. According to Gestalt theory [78], most acts dynamically unfold by an
925 internally determined organization, and these forces originate from inside. Like a classic example
926 of "soap bubble" [2], it starts as a deformed film and then reacts to local forces via soap particles
927 according to the action law, thus rendering the whole system to become a spheroid. The final form
928 of the soap is the result of dynamic self-distribution, not of artificial constraints, which guarantees
929 the thickness of the film uniformly distributed throughout the entire structure. The dynamics of
930

932 forming an equilibrium driven by internal mutual forces is well manifested in the note-taking
933 behavior on Bilibili.

934 It is most clearly seen in how participants spontaneously engaged in collective note-taking
935 on Bilibili, without external agencies in dividing and organizing the work. The majority of users
936 initially noticed the BNNote feature and were motivated to take notes through seeing other members'
937 notes shared on the platform, not through the platform's official notice. Throughout the note-taking
938 process, participants naturally distributed their efforts in contributing original notes, utilizing their
939 peer power to commit note-taking, improving on their notes against others, or directly making use
940 of others' notes, as such to ensure each video that is worth note-taking has good notes associated
941 with it in a self-organized manner. This dynamic process is not unlike the example of "soap bubble":
942 while the initial change might be triggered by some external forces, such as the introduction of
943 BNNote or the incentives brought by the clock-in activity, it is the local forces coming from the
944 community inside that cause the largely uniform distribution of quality notes throughout the
945 platform and form an equilibrium. In short, note-taking on Bilibili is such a self-organization
946 process dynamically driven by internal forces, with external factors as triggers at best, in which
947 members spontaneously act as a community whole to carry out note taking associated with videos.
948

949 6.3 Design Implications

950 Our findings show that Bilibili members make notes spontaneously as a whole community driven
951 by internal forces, which suggests several implications for design to support such collective efforts
952 "within" the community.

953 What we see here is that on a video sharing platform such as Bilibili where no instructors or
954 other facilitators are involved, it is especially important to have communal support for members
955 to mutually impel and motivate each other. As such, it becomes important to make their learning
956 activities more visible [25, 52, 70], as a way for peers to mutually motivate and hold accountable to
957 each other, which is the key for effective learning on social media platforms. Besides making notes
958 visible, we could further enhance social presence. For example, we can disclose and make more
959 visible members' note-taking efforts and their contribution to the community, such as "how many
960 notes you have taken", "how many people have referred to this note", or "how many people have
961 recommended this note", and so on. This way, people could better track their efforts, be made more
962 aware of each other's work, receive digital feedback for their efforts, and thus get more motivated
963 and more engaged in learning. When new features are introduced, we can also leverage the
964 community approach to help adopt the new feature, e.g., asking some popular community members
965 such as video uploaders to try out a new learning feature, rather than the official announcement,
966 so as to quickly make people aware and adopt it.

967 Meanwhile, it is also important to support a co-creation process where members could mutually
968 help, complement, and build on each other's work. For example, we can allow members to indicate
969 which videos they want notes but probably do not have time to do yet. We can also make members
970 easily see what notes will be highly wanted in their area, in the form of dashboard or others, so that
971 they can probably contribute ones when they have time available. Apart from sharing notes, we can
972 offer several closer collaborative mechanisms such as opening the permission for users to co-edit a
973 note, and also show how many people have contributed to a note in the form of acknowledgment
974 so people will be appropriately credited for their contribution. To enable users to collaboratively
975 complete notes for desired videos in more organized ways, we may allow members to only finish
976 part of the notes, and highlight what is being taken and what is not, so other members can easily
977 see what parts are missing and thus help fill them. Or a division of labor among members for
978 note-taking can be supported, especially for multiple episode videos, to reduce their repetitive work
979 spontaneously. We can also leverage the social approaches to allow members to learn from others'
980

notes in a cost-effective way, such as sorting all shared notes in the video’s note list by the frequency to be referred/recommended, not just by the release time. By this way, it can reduce the time cost spent by peers to seek high-quality notes as the templates for improving their note-taking skills. We can also promote the availability of online note-oriented communication, e.g., encouraging active and timely peer responses to others’ notes to further encourage each other for contribution and active learning.

In addition, based on the understanding of the complementary relationship between notes and videos, we can also consider strengths and weaknesses of these two media and the different role they play to help improve related processes. First of all, we can consider ways to reduce the cost of note-taking in such a context. For example, we can reduce the workload of extracting content within a video, such as automatically supplying video transcripts to quickly extract video structural elements. We can also leverage notes to search for videos as learning materials. For example, we can support matching note contents in addition to conventional video metadata such as titles, tags, and profiles, as the note can better reflect the overall structure and richer content of the video. To better combine the benefits of videos and notes, we can adopt appropriate forms similar to annotation-based navigation [16] or Danmu/time-anchored comments [15, 43, 85] to impose notes on some videos, thereby helping people to make full use of video’s richness and note’s abstraction in a more integrated way.

We believe, through these ways of design, members can more proactively motivate and learn from each other, more collaboratively and collectively carry out their work together, and the strengths of videos and notes could be better combined to more effectively support social-oriented VBL in video sharing contexts.

7 LIMITATIONS AND FUTURE WORK

While we reported our findings from the interview with a variety of participants who took and shared notes on Bilibili, this study has several limitations.

First, the vast majority of participants were young students or recent graduates. While young students or recent graduates were the main user population on Bilibili, it would still be valuable to cover population of other ages, occupations, and influence. Second, according to our sampling strategy, we could only find participants who had experience sharing notes. We were unable to locate those who just read notes but never took one and those who had taken notes but never shared them on Bilibili, and their note-taking practice might significantly differ from our samples. However, we could still get a glimpse of those who made notes only once while reading others’ notes multiple times and those who took many notes but shared one only. We believed these participants could fill our samples to some extent. Also, as our sampling method was that participants volunteered for the interview, our samples were likely more active than the general user community. Finally, our interviews with participants who took notes on Bilibili occurred during the early stage (less than 6 months) of the BNNote feature’s official deployment. As a result, the number of notes made by participants was relatively small. Within a short duration, users might not have developed regular note-taking practices on Bilibili. In the future, it would be helpful to further inquiry participants who have taken notes for longer periods of time (e.g., more than 6 months).

All of the aforementioned limitations may produce biased samples. Thus, interview studies with more varied samples or large-scale surveys will be expected to avoid these limitations, help corroborate our findings, and further generalize them to all video sharing platforms not specialized for learning. To get a more holistic understanding of digital note practice on Bilibili and social media in general, future work could investigate the note-taking practice across diverse age ranges, social status, and practice degrees.

1030 8 CONCLUSIONS

1031 In recent years, user-generated video sharing platforms have been increasingly leveraged to support
 1032 video-based learning, and the note-taking feature is one such example that makes the platform
 1033 support for this learning practice more explicit. In this paper, we are interested in what role note
 1034 taking plays for VBL in such a video sharing environment. To this goal, we conduct a qualitative
 1035 study based on Bilibili, which recently launched its note-taking feature called BNNote. The finding
 1036 reveals that note taking is leveraged to support a more structured and organized form of social
 1037 learning on video sharing platforms, by structuring, supplementing, and substituting the video as
 1038 well as the social processes collectively carried out by the community as a whole. We highlight the
 1039 broader roles of note taking plays for VBL beyond cognitive effects on video sharing platforms—
 1040 outside traditional, more formalized, and specialized VBL settings, e.g., making the user-generated
 1041 videos more structured, and scaffolding the social forces for more disciplined and organized learning.
 1042 On this basis, we further understand such spontaneous, dynamic, and self-organized nature of
 1043 note-taking practice on social media. We end by providing several implications to better support
 1044 social-oriented VBL, through strengthening communal ways of motivation, supporting co-creating
 1045 notes, and leveraging combined benefits of videos and notes. Through these means, we hope we can
 1046 maximize the potential of user-generated video sharing platforms such as Bilibili for social-oriented
 1047 VBL.

1048 REFERENCES

- 1049 [1] Madden Amy, Ruthven Ian, and McMenemy David. 2013. A classification scheme for content analyses of YouTube
 1050 video comments. *Journal of Documentation* 69, 5 (2013), 693–714.
- [2] Fred Attneave. 1982. Prägnanz and soap bubble systems: A theoretical exploration. *Organization and representation in perception* (1982), 11–29.
- [3] Albert Bandura. 1971. Social learning theory. Morristown.
- [4] Aaron Bauer and Kenneth Koedinger. 2006. Evaluating the effect of technology on note-taking and learning. In *Extended Abstracts of the 2006 CHI Conference on Human Factors in Computing Systems*. ACM, Montréal Québec Canada, 520–525.
- [5] Aaron Bauer and Kenneth R. Koedinger. 2007. Selection-based note-taking applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, San Jose California, 981–990.
- [6] bilibili.com. 2021. Column Angel. <https://www.bilibili.com/read/cv10976550>
- [7] bilibili.com. 2022. Financial Report. <https://www.bilibili.com/read/cv20123654>
- [8] bilibili.com. 2022. User Profile Analysis. <https://www.bilibili.com/read/cv18496411>
- [9] Robert A Bjork, John Dunlosky, and Nate Kornell. 2013. Self-regulated learning: Beliefs, techniques, and illusions. *Annual review of psychology* 64 (2013), 417–444.
- [10] Dung C. Bui, Joel Myerson, and Sandra Hale. 2013. Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology* 105, 2 (2013), 299–309.
- [11] Marina Fernández Camporro and Nicolai Marquardt. 2020. Live Sketchnoting Across Platforms: Exploring the Potential and Limitations of Analogue and Digital Tools. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12.
- [12] Yining Cao, Hariharan Subramonyam, and Eytan Adar. 2022. VideoSticker: A Tool for Active Viewing and Visual Note-taking from Videos. In *27th International Conference on Intelligent User Interfaces*. 672–690.
- [13] CCTV.com. 2019. Do you know that this generation of young people would love to study on Bilibili. <http://news.cctv.com/2019/04/17/ARTIkdxgldxCuSmVdTOimrAw190417.shtml>
- [14] Hanna Celina, Ahmed Kharrufa, Anne Preston, Rob Comber, and Patrick Olivier. 2016. SOLE meets MOOC: designing infrastructure for online self-organised learning with a social mission. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 484–496.
- [15] He Changyang, He Lu, Lu Tun, and Li Bo. 2021. Beyond Entertainment: Unpacking Danmaku and Comments' Role of Information Sharing and Sentiment Expression in Online Crisis Videos. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–27.
- [16] Mohamed Amine Chatti, Momchil Marinov, Oleksandr Sabov, Ridho Laksono, Zuhra Sofyan, Ahmed Mohamed Fahmy Yousef, and Ulrik Schroeder. 2016. Video annotation and analytics in CourseMapper. *Smart Learning Environments* 3, 10 (2016).

- 1079 [17] Si Chen, Dennis Wang, and Yun Huang. 2021. Exploring the Complementary Features of Audio and Text Notes
1080 for Video-based Learning in Mobile Settings. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in*
1081 *Computing Systems*. ACM, Yokohama Japan, 1–7.
- 1082 [18] ciavc.com. 2021. the 9th China Internet Audio & Video Convention. <https://www.ciavc.com>
- 1083 [19] M Colasante. 2011. Using Video Annotation to Reflect on and Evaluate Physical Education Pre-Service Teaching
1084 Practice. *Australasian Journal of Educational Technology* 27, 1 (2011), 66–88.
- 1085 [20] Suman Deb and Paritosh Bhattacharya. 2017. Enhancing Personalized Learning with Interactive Note Taking on Video
1086 Lectures—An Analysis of Effective HCI Design on Learning Outcome. In *International Conference on Next Generation*
1087 *Computing Technologies*. Springer, Singapore, Dehradun, India, 558–567.
- 1088 [21] Thaleia Deniozou, Mariza Dima, and Chris Cox. 2020. Designing a Game to Help Higher Education Students Develop
1089 Their Note-Taking Skills. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM,
1090 Virtual Event Canada, 181–192.
- 1091 [22] Xianghua Ding, Yubo Kou, Yiwen Xu, and Peng Zhang. 2022. "As Uploaders, We Have the Responsibility": Individualized
1092 Professionalization of Bilibili Uploaders. In *CHI Conference on Human Factors in Computing Systems*. 1–14.
- 1093 [23] Marc Exposito, Vicky Zeamer, and Pattie Maes. 2017. Unobtrusive Note Taking: Enriching Digital Interpersonal
1094 Interactions Using Gestures. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and*
1095 *Social Computing*. 167–170.
- 1096 [24] Jingchao Fang, Yanhao Wang, Chi-Lan Yang, Ching Liu, and Hao-Chuan Wang. 2022. Understanding the Effects of
1097 Structured Note-taking Systems for Video-based Learners in Individual and Social Learning Contexts. *Proceedings of*
1098 *the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–21.
- 1099 [25] Jingchao Fang, Yanhao Wang, Chi-Lan Yang, and Hao-Chuan Wang. 2021. NoteCoStruct: Powering Online Learners
1100 with Socially Scaffolded Note Taking and Sharing. In *Extended Abstracts of the 2021 CHI Conference on Human Factors*
1101 *in Computing Systems*. ACM, Yokohama Japan, 1–5.
- 1102 [26] Jennifer L Faust and Donald R Paulson. 1998. Active learning in the college classroom. *Journal on excellence in college*
1103 *teaching* 9, 2 (1998), 3–24.
- 1104 [27] Xin Fu, John C. Schaefer, Gary Marchionini, and Xiangming Mu. 2006. Video annotation in a learning environment. In
1105 *Proceedings of the American Society for Information Science and Technology*. Austin, 1–22.
- 1106 [28] Dilrukshi Gamage. 2021. Scaffolding Social Presence in MOOCs. In *Asian CHI Symposium 2021*. 140–145.
- 1107 [29] Wang Ge and Zhang Yanxiang. 2021. From "study with me" to study with you: how activities of Study With Me
1108 livestream on Bilibili facilitate SRL community. *arXiv preprint arXiv:2108.00637* (2021).
- 1109 [30] Pascal Goffin, Wesley Willett, and Petra Isenberg. 2015. Sharing Information from Personal Digital Notes using Word-
1110 Scale Visualizations. In *Proceedings of the IEEE VIS Workshop on Personal Visualization: Exploring Data in Everyday Life*.
1111 Los Alamitos, CA, United States.
- 1112 [31] Jason A Grahame. 2016. Digital note-taking: Discussion of evidence and best practices. *The Journal of Physician*
1113 *Assistant Education* 27, 1 (2016), 47–50.
- 1114 [32] James Hartley and Susan Marshall. 1974. On notes and note-taking. *Higher Education Quarterly* 28, 2 (1974), 225–235.
- 1115 [33] Nishioka Hiromi. 2021. My Korean language teachers are YouTubers: learning Korean via self-instruction. *Computer*
1116 *Assisted Language Learning* (Jul 2021), 1–30.
- 1117 [34] Matthew Kam, Jingtao Wang, Alastair Iles, Eric Tse, Jane Chiu, Daniel Glaser, Orna Tarshish, and John Canny. 2005.
1118 Livenotes: a system for cooperative and augmented note-taking in lectures. In *Proceedings of the SIGCHI Conference on*
1119 *Human Factors in Computing Systems*. ACM, Portland, Oregon, USA, 531–540.
- 1120 [35] Anam Ahmad Khan, Sadia Nawaz, Joshua Newn, Ryan M Kelly, Jason M Lodge, James Bailey, and Eduardo Veloso.
1121 2022. To type or to speak? The effect of input modality on text understanding during note-taking. In *CHI Conference*
1122 *on Human Factors in Computing Systems*. 1–15.
- 1123 [36] Khushboo Khurana and MB Chandak. 2013. Study of various video annotation techniques. *International Journal of*
1124 *Advanced Research in Computer and Communication Engineering* 2, 1 (2013), 909–914.
- 1125 [37] Kenneth A. Kiewra. 1989. A review of note-taking: The encoding storage paradigm and beyond. *Educational Psychology*
1126 *Review* 1, 2 (1989), 147–172.
- 1127 [38] Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Data-
1128 driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM*
1129 *symposium on User interface software and technology*. ACM, Honolulu Hawaii USA, 563–572.
- 1130 [39] Alexandros Kleftodimos and Georgios Evangelidis. 2016. Using open source technologies and open internet resources
1131 for building an interactive video based learning environment that supports learning analytics. *Smart Learning*
1132 *Environments* 3, 1 (2016), 1–23.
- 1133 [40] Sean Kross, Eszter Hargittai, and Elissa M Redmiles. 2021. Characterizing the online learning landscape: what and
1134 how people learn online. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–19.
- 1135 [41] Patricia G Lange. 2019. Informal learning on YouTube. *The international encyclopedia of media literacy* (2019), 1–11.

- 1128 [42] Chei Sian Lee, Hamzah Osop, Dion Hoe-Lian Goh, and Gani Kelni. 2017. Making sense of comments on YouTube
1129 educational videos: a self-directed learning perspective. *Online Information Review* 41, 5 (2017), 611–625.
- 1130 [43] Yi-Chieh Lee, Wen-Chieh Lin, Fu-Yin Cherng, Hao-Chuan Wang, Ching-Ying Sung, and Jung-Tai King. 2015. Using
1131 time-anchored peer comments to enhance social interaction in online educational videos. In *Proceedings of the 33rd
1132 annual ACM conference on human factors in computing systems*. 689–698.
- 1133 [44] Hongli Li, Min Kyu Kim, and Yao Xiong. 2020. Individual learning vs. interactive learning: A cognitive diagnostic
1134 analysis of MOOC students' learning behaviors. *American Journal of Distance Education* 34, 2 (2020), 121–136.
- 1135 [45] Min Lin, Wayne G. Lutters, and Tina S. Kim. 2004. Understanding the micronote lifecycle: improving mobile support
1136 for informal note taking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM,
1137 Vienna Austria, 687–694.
- 1138 [46] Richard K Lowe. 1999. Extracting information from an animation during complex visual learning. *European journal of
1139 psychology of education* 14, 2 (1999), 225–244.
- 1140 [47] Jia Adele Lu, Shen Siqi, Chen Shengling, Li Dongsheng, and Iosup Alexandru. 2017. An analysis on a YouTube-like UGC
1141 site with enhanced social features. In *Proceedings of the 26th International Conference on World Wide Web Companion*.
1142 International World Wide Web Conferences Steering Committee, Perth, Australia, 1477–1483.
- 1143 [48] Adele Lu Jia, Xiaoxue Shen, Siqi Shen, Yongquan Fu, and Liwen Peng. 2019. User donations in a user generated video
1144 system. In *Companion Proceedings of The 2019 World Wide Web Conference*. 1055–1062.
- 1145 [49] Nathan Magyar, Xuenan Xu, and Molly Maher. 2020. Creating and Evaluating a Goal Setting Prototype for MOOCs. In
1146 *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
- 1147 [50] K. Malterud, V. D. Siersma, and A. D Guassora. 2016. Sample Size in Qualitative Interview Studies: Guided by
1148 Information Power. *Qualitative Health Research* 26, 13 (2016), 1753–1760.
- 1149 [51] Catherine C. Marshall and A. J. Bernstein Brush. 2004. Exploring the relationship between personal and public
1150 annotations. In *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*. ACM, Tuscon AZ USA,
1151 349–357.
- 1152 [52] David W McMillan and David M Chavis. 1986. Sense of community: A definition and theory. *Journal of community
1153 psychology* 14, 1 (1986), 6–23.
- 1154 [53] Gamorán Sherin Miriam and Van Es Elizabeth A. 2009. Instructional video in e-learning: Assessing the impact of
1155 interactive video on learning effectiveness. *Journal of Teacher Education* 60, 1 (2009), 20–37.
- 1156 [54] Negin Mirriahi, Srećko Joksimović, Dragan Gašević, and Shane Dawson. 2018. Effects of instructional conditions and
1157 experience on student reflection: a video annotation study. *Higher Education Research & Development* 37, 6 (2018),
1158 1245–1259.
- 1159 [55] Toni-Jan Keith Monserrat, Shengdong Zhao, Kevin McGee, and Anshul Vikram Pandey. 2013. NoteVideo: facilitating
1160 navigation of blackboard-style lecture videos. In *Extended Abstracts of the 2013 CHI Conference on Human Factors in
1161 Computing Systems*. ACM, Paris France, 2897–2898.
- 1162 [56] Toni-Jan Keith Palma Monserrat, Yawen Li, Shengdong Zhao, and Xiang Cao. 2014. L. IVE: an integrated interactive
1163 video-based learning environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
1164 3399–3402.
- 1165 [57] Xiangming Mu. 2010. Towards effective video annotation: An approach to automatically link notes with video content.
1166 *Computers & Education* 55, 4 (2010), 1752–1763.
- 1167 [58] Pam A Mueller and Daniel M Oppenheimer. 2014. The pen is mightier than the keyboard advantages of longhand
1168 over laptop note taking. *Psychological Science* 25, 6 (2014), 1159–1168.
- 1169 [59] Pam A Mueller and Daniel M Oppenheimer. 2016. Technology and note-taking in the classroom, boardroom, hospital
1170 room, and courtroom. *Trends in Neuroscience and Education* 5, 3 (2016), 139–145.
- 1171 [60] Minoru Nakayama, Kouichi Matsuura, and Hiroh Yamamoto. 2017. Effectiveness of student's note-taking activities
1172 and characteristics of their learning performance in two types of online learning. *International Journal of Distance
1173 Education Technologies (IJDET)* 15, 3 (2017), 47–64.
- 1174 [61] omnicroreagency.com. 2022. YouTube Statistics on the Web, 2022. <https://www.omnicroreagency.com/youtube-statistics>
- 1175 [62] Ernesto Pacheco-Velazquez. 2020. Using Gamification to Develop Self-Directed Learning. In *Proceedings of the 2020
1176 International Conference on Education Development and Studies*. 1–5.
- 1177 [63] Abelardo Pardo, Negin Mirriahi, Shane Dawson, Yu Zhao, An Zhao, and Dragan Gašević. 2015. Identifying learning
1178 strategies associated with active use of video annotation software. In *In Proceedings of the Fifth International Conference
1179 on Learning Analytics And Knowledge*. ACM, Poughkeepsie New York, 255–259.
- 1180 [64] Sonny Rosenthal. 2018. Motivations to seek science videos on YouTube: Free-choice learning in a connected society.
1181 *International Journal of Science Education, Part B* 8, 1 (2018), 22–39.
- 1182 [65] Marija Sablić, Ana Miroslavljević, and Alma Škugor. 2021. Video-Based Learning (VBL)—Past, Present and Future: an
1183 Overview of the Research Published from 2008 to 2019. *Technology, Knowledge and Learning* 26 (2021), 1061–1077.

- 1177 [66] Abdulhadi Shoufan. 2019. Estimating the cognitive value of YouTube's educational videos: A learning analytics
1178 approach. *Computers in Human Behavior* 92 (2019), 450–458.
- 1179 [67] Franck Silvestre, Philippe Vidal, and Julien Broisin. 2014. Tsaap-Notes—An Open Micro-blogging Tool for Collaborative
1180 Notetaking during Face-to-Face Lectures. In *2014 IEEE 14th International Conference on Advanced Learning Technologies*.
1181 IEEE, 39–43.
- 1182 [68] sohu.com. 2017. Note-taking APP: moving toward management and community. https://www.sohu.com/a/147366996_783965
- 1183 [69] Anselm Strauss and Juliet Corbin. 1998. *Basics of qualitative research: Techniques and procedures for developing grounded*
1184 *theory*. Sage Publications, Thousand Oaks, CA.
- 1185 [70] Na Sun, Mary Beth Rosson, and John M Carroll. 2018. Where is Community Among Online Learners? Identity, Efficacy
1186 and Personal Ties. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. ACM, Montreal
QC Canada, 1–13.
- 1187 [71] Loeng Svein. 2020. Self-directed learning: A core concept in adult education. *Education Research International* 2020
1188 (2020).
- 1189 [72] Chien-Lin Tang, Jingxian Liao, Hao-Chuan Wang, Ching-Ying Sung, and Wen-Chieh Lin. 2021. ConceptGuide:
1190 Supporting Online Video Learning with Concept Map-based Recommendation of Learning Path. In *Proceedings of the Web Conference 2021*. ACM, Ljubljana Slovenia, 2757–2768.
- 1191 [73] tech.huanqiu.com. 2014. Net Ease CEO Ding Lei: digital note taking is likely to become the next big Internet-based
1192 service. <https://tech.huanqiu.com/article/9CaKrnJB3HV>
- 1193 [74] T. Valtonen, S. Havu-Nuutinen, P. Dillon, and M. Vesisenaho. 2011. Facilitating collaboration in lecture-based learning
1194 through shared notes using wireless technologies. *Journal of Computer Assisted Learning* 27, 6 (2011), 575–586.
- 1195 [75] Julita Vassileva. 2008. Toward social learning environments. *IEEE transactions on learning technologies* 1, 4 (2008),
199–214.
- 1196 [76] Lev Semenovich Vygotsky and Michael Cole. 1978. *Mind in society: Development of higher psychological processes*.
1197 Harvard university press.
- 1198 [77] Hung-chun Wang and Cheryl Wei-yu Chen. 2020. Learning English from YouTubers: English L2 learners' self-regulated
1199 language learning on YouTube. *Innovation in Language Learning and Teaching* 14, 4 (2020), 333–346.
- 1200 [78] Michael Wertheimer. 2010. *A Gestalt Perspective on the Psychology of Thinking*. Towards a Theory of Thinking. Berlin,
Heidelberg.
- 1201 [79] Wesley Willett, Pascal Goffin, and Petra Isenberg. 2015. Understanding Digital Note-Taking Practice for Visualization.
1202 *IEEE Computer Graphics and Applications* 35, 4 (2015), 38–51.
- 1203 [80] Amber E Witherby and Sarah K Tauber. 2019. The current status of students' note-taking: Why and how do students
1204 take notes? *Journal of Applied Research in Memory and Cognition* 8, 2 (2019), 139–153.
- 1205 [81] Sarah Shi Hui Wong and Stephen Wee Hun Lim. 2021. Take notes, not photos: Mind-wandering mediates the impact
1206 of note-taking strategies on video-recorded lecture learning performance. *Journal of Experimental Psychology: Applied*
(2021).
- 1207 [82] Qunfang Wu, Yisi Sang, and Yun Huang. 2019. Danmaku: a new paradigm of social interaction via online videos. *ACM*
1208 *Transactions on Social Computing* 2, 2 (2019), 1–24.
- 1209 [83] Xiang Xiao and Jingtao Wang. 2017. Understanding and detecting divided attention in mobile MOOC learning. In
1210 *Proceedings of the 2017 CHI conference on human factors in computing systems*. 2411–2415.
- 1211 [84] Yaxing Yao, Jennifer Bort, and Yun Huang. 2017. Understanding Danmaku's potential in online video learning. In
1212 *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*. 3034–3040.
- 1213 [85] Matin Yarmand, Dongwook Yoon, Samuel Dodson, Ido Roll, and Sidney S Fels. 2019. "Can you believe [1: 21]?" Content
1214 and Time-Based Reference Patterns in Video Comments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–12.
- 1215 [86] Lee Yoonjoo, Chung John Joon Young, Song Jean Y, Chang Minsuk, and Kim Juho. 2021. Personalizing Ambience
1216 and Illusionary Presence: How People Use "Study with me" Videos to Create Effective Studying Environments. In
1217 *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13.
- 1218 [87] A M F Yousef, M A Chatti, N Danoyan, H Thüs, and U Schroeder. 2015. Video-Mapper: A Video Annotation Tool to
1219 Support Collaborative Learning in MOOCs. In *Third European MOOCs*. Mons, Belgium, 131–140.
- 1220 [88] Ahmed Mohamed Fahmy Yousef, Mohamed Amine Chatti, and Ulrik Schroeder. 2014. The State of Video-Based
1221 Learning: A Review and Future Perspectives. *International Journal on Advances in Life Sciences* 6, 3 & 4 (2014), 122–135.
- 1222 [89] Carol Zander, Lynda Thomas, Jan Erik Moström, and Anna Eckerdal. 2020. Copying Can Be Good: How Students
1223 View Imitation as a Tool in Learning to Program. In *2020 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.
- 1224 [90] Dongsong Zhang, Lina Zhoua, Robert O. Briggs, and Jay F. Nunamaker Jr. 2006. Instructional video in e-learning:
1225 Assessing the impact of interactive video on learning effectiveness. *Information & Management* 43, 1 (2006), 15–27.

- 1226 [91] Ying Zhang, Luming Zhang, and Roger Zimmermann. 2015. Aesthetics-guided summarization from multiple user
1227 generated videos. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 11, 2 (2015),
1228 1–23.
- 1229 [92] Rebecca Zheng, Marina Fernández Camporro, Hugo Romat, Nathalie Henry Riche, Benjamin Bach, Fanny Chevalier,
1230 Ken Hinckley, and Nicolai Marquardt. 2021. Sketchnote Components, Design Space Dimensions, and Strategies for
1231 Effective Visual Note Taking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM,
1232 Yokohama Japan, 1–15.
- 1233 [93] Meina Zhu. 2022. Designing and delivering MOOCs to motivate participants for self-directed learning. *Open Learning: The Journal of Open, Distance and e-Learning* (2022), 1–20.
- 1234 [94] Barry J Zimmerman and Dale H Schunk. 2011. *Handbook of self-regulation of learning and performance*. Routledge/Taylor & Francis Group.
- 1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274