THE CHINESE EXAMINATION SYSTEM IN DYNASTIC CHINA: DID IT SELECT THE BRIGHTEST AND BEST?

DAVID L. MCMULLEN
Department of East Asian Studies
Faculty of Asian and Middle Eastern Studies
University of Cambridge
dlm25@cam.ac.uk

ABSTRACT
This article describes the principles behind the Chinese examination system in dynastic China, outlines some of the changes that took place in the examinations over more than a millennium and provides a tentative answer to the question of how effective they were as selectors for government service of the brightest and best from among China's population.

INTRODUCTION
Western assessment of the Chinese heritage and Chinese history is in a state of constant change. There have been times when the European intellectual community admired China greatly, and times when opinion was very much more negative. Broadly, the West has admired a strong China and tended to denigrate a weak China. Thus in the eighteenth century China was admired, while in the nineteenth and into the twentieth, opinion was generally much more qualified. Now that China is stronger again, there is another shift in perspective. That shift has not yet fully been played out.

There have, however, been two features of Chinese history for which estimates have been relatively constant: first, the contributions of Chinese technology have been acknowledged. That is not solely due to the work of the Cambridge scientist and historian Joseph Needham, but rather has run through Western assessment of Chinese culture from the eighteenth century on.

There is, secondly, another, and rather unusual, aspect of the Chinese heritage that has always been admired, through good and bad times, and that is the public examination system for recruitment to the civil administration of the Chinese empire. It is unusual, because admiration for the Chinese heritage has usually been for artistic and technological rather than institutional achievements. The immensely influential compilation on China edited by du Halde and first published in English in 1736, opened the third volume of its comprehensive account with a description of Chinese education and of the public examination system. It brought out a number of features of the system, including the strict security surrounding examinations and the standardization of the system throughout the empire. Its criticisms were only that it served to identify the elite in a way that was not so in Europe and that its syllabus was relatively narrow.
EUROPEAN ADMIRATION

Even when estimates of China were low, as for example in the aftermath of the Opium War, Western observers admired the examination system. T.T. Meadows (1847), a shrewd and sober observer of the mid-nineteenth century scene, attributed the longevity of the Chinese empire solely to its implementation of an effectively meritocratic public examination system:

“It is by constantly enlisting on their behalf faith in a uniform policy of which no other nation furnishes an example, men from all classes, with talented and determined minds…that the founders of successive dynasties have established themselves on the throne of China…”

“It is, then, to the exclusive advancement of merit and talent, quite independent of every other principle or doctrine, that the Chinese empire is indebted for its long duration.”

That is another way of saying that the impressive continuity in Chinese institutional history, running through successive dynasties, owes to its ability to recruit public servants from a broad social base and on the basis of tested merit. While perhaps few would make the point as emphatically as Meadows, most would not deny that the system greatly contributed to the stability of successive regimes. It is generally accepted, also, that the adoption of public examinations for entry into the civil service, first in Great Britain in the late nineteenth century, then in France and the U.S.A., was the result of the prestige and influence of the Chinese model. In other words, this was a case—a rare case—of direct institutional influence of an admired feature of the Chinese administrative model.

This respect for an institutional tradition of great duration and sophistication points up the question; how effective and important was the system during the different periods in Chinese history that it operated?

THE PREHISTORY OF THE EXAMINATIONS

The obvious and remarkable thing about the examinations is that they make the meritocratic principle unambiguously central to one of the most important operations that any advanced human society conducts, the administration of the state. They privileged worth over birth. I say one of the most important, but in fact, to many, it was THE most important in China, apart from being the emperor himself. The official hierarchy, the relatively tiny elite who formed the apex of the administrative hierarchy governing China’s vast territory, carried immense prestige. There was simply no other career that could come anywhere near rivalling that of government official. Access to this hierarchy, in the late imperial period at least, was through the public examination system.

It is as well to bear in mind that this meritocratic principle has a very long history in China, one that far precedes its institutionalization in the examinations themselves. In the Confucian Analects, Confucius is recorded as having emphasized that recruitment for official service should depend on moral character and abilities, and that there should be a process of selection involved. Following this, almost all the main schools of statecraft in the

---

2 Meadows, p. 152.
pre-Qin period, but especially those that contributed most to the Chinese ideology of the unitary state in China, the Confucian and Legalist schools, advocated selection for administrative service on the basis of worth.

**EVOLUTION IN PRACTICE**

A summary of the long process by which this principle of selection on the basis of merit became institutionalized shows two things. The first is that from the Han dynasty on, until the abolition of the civil service examinations in 1905, the system evolved enormously. That is to say that features of the examinations conducted in the eighth century would barely be recognizable to candidates who sat them in the nineteenth. However, much of the terminology and a number of the ritual procedures remained constant. For, ‘Tang precedents were extraordinarily influential in later times.’ For example the *jinshi*, first introduced in the Sui dynasty (581 – 618) attained great prestige by the late seventh century and remained as the most prestigious of all examinations until the system was abolished. Also, the final examinations conducted at the capital were the responsibility of the Board of Rites, from 737 on. The Board still conducted the examinations when they were abolished nearly twelve centuries later. Certain features of the ceremonial or ritual procedures surrounding the examinations also remained constant. A celebratory feast and expression of thanks to the chief examiner was one of these.

One of the features that changed most concerned the relationship of candidates to the examiners. In the early period of the examinations, candidates were expected to make themselves known to examiners and to canvas them vigorously. They were expected to promote their own candidacy, by showing their literary compositions and by advertising their social and moral standing. Examiners even complained about the mountains of compositions that candidates had sent them before the examinations. This feature of the Tang system would seem to run counter to the ideal implicit in meritocracy. Indeed, the principle of the ‘anonymity of the candidate’ (*hu ming*) was recognized, but by no means accepted.

The Song dynasty, following on after the Five Dynasties period when China was not subject to a unitary government, has been called a period of ‘precocious modernity’. The Song may have been militarily weak; but culturally it was immensely sophisticated. Its bureaucratic structure was more developed than any before it in China. In its implementation of the examination system, it certainly followed on from the Tang. But it tightened up its procedures to a spectacular degree. One of its principles was precisely the implementation of the ‘anonymity of the candidate’. No examiner was allowed to know whose scripts he was assessing. To enforce this, the administrators had all scripts copied by professional clerks. Each script was copied twice. The two scripts were then independently assessed and the results of the two examiners sealed independently. The lists were then integrated by a third examiner. This process must surely have helped to select the brightest, on the basis of their performance, not their status or connections. The principle of the anonymity of the candidate was implemented by later dynasties, although elaborate ways of circumventing it grew up. In other respects, the Song system is a half-way point between the Tang and the system as it stabilized in the late imperial period, from the Ming right through until the system was abolished.
A second difference was that in the first centuries of its existence, the system selected only a small proportion of the body of officials. In effect, it identified not the entire official body as in later periods, but rather the elite within the civil administration, the ‘fast stream’, the future close servants \((jinchên)\) of the emperor, whose prestige was immense. One of the main reasons for its eventual triumph as the only means of access to officialdom was undoubtedly its success in identifying political leaders in this way. Over the Tang, the numbers of chief ministers who were graduates climbed steadily, until it reached a point where a \(jinshi\) degree was almost a required qualification for the highest offices. After that, it was a relatively straightforward step to increase the intake of officials admitted through examination success, until by Ming times it had indeed become the universal point of entry.

This issue of the evolution of the system is relevant to the title for two other, related reasons: first, institutional change in China always provoked comment and criticism. In the case of the examination system, so many were involved and so deeply did the outcomes affect them and their communities that the evidence for attitudes over twelve centuries is very copious. There is both critical comment and an enormously rich tradition of anecdote, both celebratory for the few who succeeded and despairing for the very large numbers destined to fail. Secondly, were the examinations an effective way of selecting the best for public service?

The question of whether it selected the brightest from among the population can be tackled from different angles: (i) from the point of view of the nature of the syllabus or the content of the examinations, and (ii) from the point of view of the conduct of the examination procedure itself.

### THE SMALL SIZE OF THE MANDARINATE

There is, however, one other point of general relevance to the topic of this paper, and that deserves emphasizing. This is that the mandarinate remained extremely small, throughout successive dynasties. In the Tang, it has been estimated at about 17,000-19,000, governing an empire with a population of about 50 million to 70 million. China’s population was, of course, never static; but it does not show a steady smooth increase curve either. Even when in Qing times, the population of China expanded greatly, to perhaps 200 million in 1700 and 300 million in about 1800, with enormous implications for the numbers of candidates, the numbers of mandarins, and the structure of provincial administration remained approximately the same at perhaps 25,000 or so. This has two quite different implications.

It meant in the first place that the officials selected through the public examinations thus represented only the highest echelons of the administrative hierarchy. Below them were the very much more numerous permanent clerical staff in any locality or agency. These clerical staff \((lixuí)\) and the yamen runners and lictors beneath them were not at all subject to the selection by examination. Moreover, in many cases their jobs were considered hereditary.

The interface between the mandarins, the elite, who were constantly rotated from post to post in the provinces, and the clerical and yamen staff, who remained forever in the same locality, was first identified as a problem in the Song dynasty. It remained a deeply entrenched feature of the system and on the agenda for reformers for ever after that. Essentially the issue was that the clerical staff controlled the mandarins, or drastically
curtailed their real power. So, if we are analysing the overall political scene in dynastic China, the meritocratic ideal, very important though it was, only related to the highest layers of the administrative pyramid. The vastly more numerous clerical staff, who played so vital a role, was, as far as the meritocratic ideal was concerned, ‘beneath the radar’, as we say nowadays. They had moral compasses that tended to face in altogether different directions.

Secondly, the small numbers of the mandarins meant that, to put it bluntly, the officials conducting the examinations were greatly overworked. There were said to be about three million candidates for shengyuan status every second year (Elman, 2000, p. 237). But the numbers of qualified officials remained practically static. This in turn had very serious implications for the efficiency and thoroughness of the examination process itself. Let us turn now to the process of assessment itself and see what problems it faced.

THE PROCESS OF THE ASSESSMENT

This situation greatly affected both the procedures governing the examinations and their content. The examination procedure itself placed an enormous burden on local officials. In the early period, there were local, that is to say county level and prefectural level tests, conducted by the officials who held office in the locality at the time. These are relatively informal, and not so much is known about them. But the questions set in some of these local examinations do survive.

The late imperial examinations involved an empire-wide operation of quite colossal scale and complexity. Only recently has scholarship come to bring out the full complexity and importance of local examinations, from county level up, in the cultural and administrative life of both officials and candidates alike. They were, literally, major local events attended by ceremony and viewed by crowds. There were three stages, required to reduce the enormous numbers of candidates. In order to answer the challenge of the topic, it is necessary to summarize the content of all three.

The most local examinations were conducted at ‘district’ or ‘county’ levels (xian). These were in a sense preliminary, because success in them merely qualified a candidate to be a ‘licentiate’. A preliminary stage involved qualifying examinations for entry into the official school system. They were held in very formal conditions at the county yamen, and comprised several stages. The syllabus required rote knowledge of the Four Books and the Five Classics, and the ability to compose shi and fu on a set theme and to prescribed rhymes. The candidates also had to prove that they knew the Sheng-lun guangxun by heart.

The next stage was the prefectural examination, and it also consisted of three parts, identical to the county examination. Following that was the provincial examination which was what has been called the ‘qualifying examination’. Again, the syllabus was on the Four Books. Success in all these examinations provided a candidate with shengyuan status.

Shengyuan were in turn required to take annual examinations, or more properly re-examinations. But their principal ambition was to sit the next tier of examinations. There was a preliminary provincial examination, again held, every third year, with much ceremony

---

3 For information on the late imperial system included in the following pages, I am greatly indebted to Benjamin Elman, A Cultural History of Civil Examinations in Late Imperial China (Berkeley: University of California Press, 2000).
and formality, in provincial capitals. Candidates were under extreme pressure and cases of break-down or illness were common. It was a tradition, for example, that the dead could take revenge for candidates’ past misdemeanours precisely at the time when they sat examinations. Again, the syllabus did not go beyond the *Four Books* or the *Five Classics*. But there was also a question on current political problems, just as there had been in the Tang period. However, these examiners deliberately avoided anything controversial, and they usually set topics concerning the past rather than the present. Candidates, moreover, wrote answers that were as inoffensive as possible. In other words, this was not a real invitation to critique current policies at all.

Finally came the metropolitan examination, the *huishi*, held every third year in Beijing. This also had three sessions. Again the syllabus did not go beyond the *Five Classics* and the *Four Books* and their commentaries. After the third stage, candidates proceeded, after further formalities, to take the palace examination. This was the culmination of the whole process, and the Emperor himself symbolically presided over it.

The number of scripts was enormous. Assessment has been called ‘a complicated guessing game’. Even highly conscientious and exemplary examiners conceded that the task of providing accurate evaluation was daunting and, fundamentally, unrealizable. To visualize this, we have to remember that the Chinese empire ran on paper. By Ming times, it had done so for at least six centuries, to a time long before paper was a common currency in Europe. It is no accident that some of the great Tang poets, like Du Fu in 758 A.D., writing at Huazhou to the east of Chang’an after he had been dismissed from his court post, complained as a humble county official of the piles of documents on his desk. Bai Juyi and others did the same. Chief examiners, of course, had assistants, even at the early stages of the examination system. But they too were overworked and subject to unrealistic deadlines.

In the early period, however, the examiners had it relatively easy, and the system might have been able to cope. In the Tang, the numbers of *jinshi* candidates were kept down by, first, the lower rates of the necessary degree of literacy, secondly by the restriction of numbers of approved candidates coming in to Chang’an from the provinces, and thirdly by the rigorous checking system enforced by the officials who vetted the candidates and permitted them to proceed for examination. Probably only about one thousand candidates sat the *jinshi* every year.

But in later periods, the numbers increased dramatically. For example, in 1523, 19 examiners had gone through some 3,600 scrolls for each of the three sessions of the examination. In 1762, for the Hunan provincial examinations, 4,000 and more candidates participated, producing 12,000 scripts, in the three sessions. This broke down to 56,000 individual answers. There was necessarily a time limit to examining - 18 days and nights. The chief examiner, the famous scholar Qian Daxin, confessed that he had no great confidence in the resulting grading (Elman, 2000, p.424).

The tendency was for examiners to simplify assessment by giving undue weight to one part of the examination, making it the determinant for success or failure. But in different periods, different parts of the examination were given priority. It is time now to turn to the content of the examinations. What sort of knowledge and what sort of skills did the system test for?
THE CONTENT OF THE EXAMINATIONS

In the Tang, the examination system taken as a whole was at its broadest and most inclusive. There were examinations in all the subjects that were taught at the Guozi jian, the state educational agency at Chang’an. These subjects included mathematics (suan 算), penal law (lu 律) and orthography (shu 書), the variant forms of the Chinese script and their history.

But there was a problem here. This was that these specialist subjects simply did not command enough prestige for them to be taken by members of the elite. They are as a consequence poorly documented, and no questions or answers survive. These specialist examinations never occupied an important place in the later system. There were other examinations, for example in medicine, that were outside the remit of the Guozi jian, and the content of which is unknown.

The content of the main examinations, instead, remained focussed on two main skills. The first was knowledge of the body of ancient texts known as the Confucian classics. The second was an ability to compose elegant, erudite but relatively short answers to topics prescribed by the examiners. These were again of two types, prose poems and answers to questions on ‘current affairs’. To perform adequately in these exercise, candidates had to have control of what has aptly been called the ‘memorization corpus’.

The one examination for which questions and answers survived in considerable answers was the jinshi. This became the most prestigious of the range of examinations towards the end of the seventh century. It went on in all the dynasties that followed to be the high point in the whole system. It had three parts: a test of rote memory of the classics; a section on discussion of contemporary affairs and a section on composing shi and fu, on set topics and to a set rhyme scheme.

Let us look briefly at the questions on contemporary affairs. The Tang has a special place in the ‘master narrative’ of Chinese history, as being a time of relative openness. Openness towards foreign influences, openness towards women, openness in politics. The questions on contemporary affairs certainly seem to reflect this. For example, a question put to candidates in 766 A.D. in the very far south of China, at the prefectural town of Daozhou on the Hunan-Guangdong border, asked for the causes of the recent tenfold rise in prices, in grain and silk. Candidates were required to comment on the change of attitude in the labour force that had resulted in a drop in production and a rise in value. The same prefect asked for an explanation for the growing provincial separatism of the period, or why it was that war-lords commanding their own armies were no longer willing to accept the authority of the court and the emperor.

Even the poet Du Fu asked questions that had bearing on the decline of central authority. This was again when he was on the staff at Huazhou, east of the capital. He invited comment on the proliferation of envoys from the capital passing through his jurisdiction to the provinces to the east, sent out in an effort by the emperor and the court to regain control of financial and military affairs after the catastrophic rebellion of An Lushan in 755 A.D. .

In the Tang there was also another category of examination that called for comment on the contemporary political situation. This was the ‘decree examination’ (zhi ke 制科), and some of the answers by candidates amounted to comprehensive reformist reviews of the political state of the dynasty. Of course, candidates knew well enough that they should not
be too keenly critical of authority. There was too a language of reformist suggestions that was traditional and therefore had the protection of custom. We have fascinating answers, for example, by the famous poet friends Yuan Zhen and Bai Juyi. But there were cases where candidates broke the rules and were too critical. One such happened in 828 A.D., and we have both a record of the enquiry that followed and the candidate’s answer. He did not manage a significant career after this.

However, the fact was that the section on the contemporary problems, fascinating although it is to a modern scholar looking at the record, was less important for examiners than the final section, requiring composition of shi and fu on set themes. In the early period, from the late seventh century on, there was a relentless trend to making the examinations, which incidentally were held at the capital every year rather than every three years as in the late imperial period, above all a test of skill in literary composition.

As a reminder, the most prestigious examination, the jinshi, consisted of three parts, a test of rote knowledge of the Confucian canonical texts; a series of questions on contemporary affairs; and a literary composition, the composition of poems or prose poems (shi or fu). The topics set were relevant to the general issue here: candidates tended to be asked to compose on what might be called the dignified rather than the efficient aspects of the dynastic state: great state buildings, the palaces, state rituals, the imperial gardens as the site of natural phenomena; the grand state rituals.

The way in which skill in literary composition came to dominate the jinshi in the eighth and ninth centuries caused a lot of dissatisfaction at the time. There were trenchant criticisms of the obsession with verbal dexterity, and concern for frivolous aesthetic that lay behind it. The standard criticism, which ran through until the end of the system was voiced in a paraphrase of an ancient legalist epigram. Students “Study things they will never use; and use things they have never studied” (Suo xi fei suo yong; suo yong fei suo xi 所習非所用所用非所習). Examination learning was called ‘empty talk on paper’ (zhi shang kong tan 紙上空談; or wu yi zhi kong yan 無益之空言).

In effect, these critics asked precisely the question that forms the title of this paper: what relevance had such skills to the government of the empire and to the solution of the grave crises that confronted the emperor and the court? There were also repeated complaints that the examinations being only paper tests, did not examine candidates’ moral character or conduct. In the Qianlong period, serious attempts were made to require moral assessments of shengyuan.

The requirement for composition remained a component of examinations throughout the late imperial period. But after the Tang and Song, the spotlight of reformist concern shifted to other components of the examinations.

Similarly, the requirement for an ability to write on contemporary affairs also remained until the Qing period. In the Song, a surprising degree of technical knowledge was required, only for the system to discover that examiners were not up to the level of knowledge that some candidates commanded. However, some knowledge of the calendar, of astronomy and of music was required. In fact, the record of what these questions contained is quite full.

Many of the questions had a historical dimension that, in effect, eliminated the possibility of controversy. Policy questions tended to be downgraded and increasingly undervalued in importance). There were calls for them to be more meaningful.
Occasionally, meaningful policy questions were asked, by reformist examiners. The famous canonical scholar Ruan Yuan (1764–1849), as Education Commissioner for Zhejiang from 1795 to 1798 for example, even asked candidates for their views on how cheating could be eliminated. Elman, 2000, p.235)

Instead of these components, the later system prioritized another one. The case of the ‘eight legged essay’ (bagu wenzhang) is perhaps the best known example. This first featured in the 1480s. It featured a very formal and disciplined internal organization. Even the lay-out of the essay in the candidates’ scripts, the actual writing on the page, had to reflect its structure. By nineteenth century times, it had come to stand for the whole examination procedure.

There were times when other parts of the procedure received undue weight. Essays on the Four Books were an example; in the late period, these received more weight, and other sections of the examination merely provided corroborative evidence. Still, the burden on overworked and exhausted examiners remained relentless. One of the themes in the popular lore that grew up round examinations was that of the carelessness of the derelict examiner confronting in the underworld the ghosts of the thousands of candidates he had failed.

CONCLUSION

The public examination system in dynastic China functioned in different ways in different periods. In the Tang, it functioned to identify the fast stream, the highly educated officials who took over the central civil posts. It is very conspicuous that over the Tang more and more of the chief ministers were examination graduates. They made their way up partly through political connections and partly because they had been able, from the start, to obtain better training and education.

We cannot possibly say that the examination system in the late medieval period, the Tang, identified the brightest and best in the population as a whole. What it did do was to identify intellectually able people, both from aristocratic families and from the second tier of elite families, and point them in the direction of successful careers. Many of them were also from old aristocratic families; the meritocratic ideal was far from universally implemented. Nonetheless, it was there.

In the Song, wealth was much more widely spread, literacy increased and the training needed to take part in the public examinations was much more widely available. An extraordinary preoccupation with integrity in the bureaucratic system and in political life resulted in a new emphasis on the meritocratic principle. Examinations were seen as a substantive test of a man’s real abilities. The number of jinshi increased fourfold. By the late eleventh century, the number of jinshi exceeded the number of posts available. Also, the school system at one stage was expanded and even threatened to displace the examinations system as a route to office. This did not happen, mainly because ultimately, the examinations system cost less.

The same dynamic, the need for recruitment of able and loyal public servants, operated at the start of the Ming and the Qing dynasties, and this ensured that the system prospered. As time went on, especially during the long pax sinica of the eighteenth century, the numbers of candidates greatly outnumbered the ultimate number of posts available.
Therefore, examiners made their main priority not the identification of great talent, but the need to fail most of the candidates who came forward. Technicalities rather than considerations of real quality were manipulated to eliminate as many candidates as possible. Without doubt, some of the very able, intellectually energetic and rigorous candidates did get through and proceed to illustrious careers.

The examinations never fully realized a principle that is probably un-realizable in any culture and any context. A number of factors operated to result in the examinations not being open to all. Being a candidate was expensive at all stages. The basic education needed was available only to well-off families, lineages or clans. The costs of being a candidate, in travel costs, living expenses, tips, fees and so on, were very considerable. This meant that, ‘..those in the competition were a self-elected minority of young men from literati or merchant families, lineages or clans, with sufficient linguistic and cultural resources to invest in their male children.’ (Elman, 2000, p.249). It has been said that the cost of taking part in the triennial examinations in Beijing was the equivalent of buying a round-the-world air ticket. ‘..the educational curriculum and its formidable linguistic requirements effectively eliminated the lower classes from the selection process.’ (Elman, 2000, p.xxiv). There were regional imbalances, since the necessary preparation was much more readily available in rich and prosperous areas in which academic traditions could be consolidated. There was also, it has to be said, a significant element of corruption, impersonation, cheating and bribery.

None the less, the figures available are impressive. There are tentative statistics available to suggest how many candidates from new backgrounds were successful in the later system. These statistics are drawn from jinshi rolls, the lists of those successful in the highest examination. Candidates who had commoner status, rather than licentiate status or higher, comprised 33% under the Song; 49.5% under the Ming and 37.6% under the Qing.

Yet the re-evaluation of the examination system made in recent decades sees it not as a genuinely impartial mechanism for throwing open access to high political power and elite status, but rather a process of inclusion and exclusion based on tests of classical literacy that effectively denied access to those who had not received the required classical training.

Success rates remained extremely low throughout the long centuries the system operated. In the Tang, the numbers of candidates for the jinshi were low, perhaps one thousand annually. The pass rate for the jinshi was between two and three per cent. In the Song, the numbers of candidates increased dramatically, and more were passed, perhaps amounting to one in fifty of all candidates. In the Ming and Qing, the success rate for shengyuan taking the jinshi was said to be one in three thousand. The system is extraordinarily well documented. We know that in the Ming and Qing periods, 51,341 candidates obtained the jinshi degree. Whether this number of successful candidates did in fact represent the brightest from among the population of China is highly dubious.

**MAIN SOURCES**


